Optimal investment under behavioural criteria in incomplete markets

José Gregorio Rodríguez Villarreal

Doctor of Philosophy
University of Edinburgh
2015

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*José Gregorio Rodríguez Villarreal*)

A Adriana, Miguel, Alma, Josefina, Ramón y Christián.
Por estar ahí...

Publications

This thesis is based upon joint work between the author and my advisor Miklos Rásonyi contained in [42], published in 'Advances of Mathematics in Finance', Banach Center Publications 104, 2015 and, a preprint, currently available at arXiv:1501.01504.pdf. The results of Chapter 3 are based on the first paper [42], and results and methods developed in Chapter 6 are drawn from the second. Chapter 4 and 5 provide preliminary results needed to setup our model and to obtain our results in Chapter 6, on the other hand Chapter 2 describes the main results applied in Chapter 3.

Acknowledgements

I would like to acknowledge first and foremost, my supervisor Miklós Rásonyi for his guidance, enthusiasm, patience and support throughout my PhD studies. For posing interesting problems and for being always willing to give sensible and valuable advice.
I would also like to acknowledge the support of the School of Mathematics for providing tools and learning opportunities as well as providing a stimulating environment to do research. I also want to express wholehearted gratitude to the Probability and Stochastic Analysis group that encouraged me through my studies and for all the learning opportunities and invitations to talks, seminars and courses.
This process would not have reached an end if I had not had the strong and tireless encouragement from Adriana, Alma and Miguel *¡Gracias por estar siempre conmigo!*; and also from (mi tío) Christian González and his family and from Dorado Santana family.
I also would like to express my gratitude to all the people in the Maths department that help me throughout my PhD studies. I would like to thank all my friends in Edinburgh and in México for their support, concern, criticism and good moments. You know who you are, *gracias*.
Finally I would like to acknowledge the financial, operational and legal support provided by the University of Edinburgh and the National Council of Science (CONACYT) in Mexico.

Lay Summary

This work is concerned with the study of optimal investment problems in Mathematical Finance. While this is a widely analysed subject in the literature under the tenets of Expected Utility Theory, and this framework has become a standard topic for economists and financial analysts, many of its principles have been subject to an intense debate. Indeed, many of the underlying principles have been questioned, thus, a more general framework that could give a satisfactory explanation of agent's choices and preferences was absolutely necessary.
We analyse the problem of optimal investment when preferences of market participants are described within the framework of Cumulative Prospect Theory (CPT), developed by Daniel Kahneman and Amos Tversky in the last two decades. We translate the problem into a mathematical optimisation problem and propose verifiable conditions in order to ensure well-posedness of the problem as well as existence of optimal strategies for discrete-time models and for a fairly broad family of models in continuous time, improving significantly past results in the field. Our results does not assume any type of completeness of the market, whatsoever.
Our aim is to have a self-contained and detailed study of the problem and of our own developments in the problem of optimal investment under general performance criteria and with straightforward assumptions. It is worth pointing out that it is possible to regard Expected Utility Theory and some of its generalisations as particular cases of the problem of optimal investment in Cumulative Prospect Theory, therefore our methods and conclusions apply to the standard theory. The results presented in this work generalize and improve significantly the theorems concerning quantitative Behavioural Finance within the framework of CPT. In addition, we think that the ideas developed here can be adapted and applied to obtain further results in problems about optimal investment as well as other applications.

Abstract

In this thesis a mathematical description and analysis of the Cumulative Prospect Theory is presented.
Conditions that ensure well-posedness of the problem are provided, as well as existence results concerning optimal policies for discrete-time incomplete market models and for a family of diffusion market models.
A brief outline of how this work is organised follows. In Chapter 2 important results on weak convergence and discrete time finance models are described, these facts form the main background to introduce in Chapter 3 the problem of optimal investment under the CPT theorem in a discrete time setting. We describe our model, present some assumptions and main results are derived. The second part of this work comprises the description of the martingale problem formulation of diffusion processes in Chapter 4. A key result on the limits and topological properties of the set of laws of a class of Itô processes is described in Chapter 5. Finally, we introduce a factor model that includes a class of stochastic volatility models, possibly with path-depending coefficients. Under this model, the problem of optimal investment with a behavioural investor is analysed and our main results on well-posedness and existence of optimal strategies are described under the framework of weak solutions.
Further research and challenges when applying the techniques developed in this work are described.
Keywords: Optimal portfolio ; Behavioural Finance ; Probability distortion ; Well-posedness ; Optimal investment ; Martingale problem.

Notation

- d or n will refer to the dimension of an Euclidean space $\mathbb{R}^d$ or $\mathbb{R}^n$.
- $intG$ the interior of a subset G.
- $\langle \cdot, \cdot \rangle$ denotes an inner product.
- $\overline{\lim} \{x_n\}$, the 'limsup' of the sequence $\{x_n\}$ defined as $\overline{\lim} \{x_n\} := \inf_n \sup_{k \geqslant n} x_k$. Similarly, $\underline{\lim} \{x_n\} := \sup_n \inf_{k \geqslant n} x_k$
- $\overline{X}$ or $cl(X)$ the topological closure of a set X.
- $\|X\|_\infty := \inf\{M > 0 : \mu(\omega \in E : |X(\omega)| > M) = 0\}$, a norm in the Banach space $L^\infty(E, \mathcal{E}, \mu)$.
- $\mathcal{L}(X) := \mathbb{P} \circ X^{-1}$, the law of a random variable X.
- $\mathcal{M}^1(E)$ the space of probability measures on $(E, \mathcal{E})$, a measurable space.
- $\mu_n \Rightarrow \mu$ denotes convergence of the sequence $\{\mu_n\}_n \subset \mathcal{M}^1(E)$, to μ in the weak topology.
- $supp(\mu)$ the support set of the measure μ.
- $S : \Omega \rightrightarrows \mathbb{R}^d$ a set-valued function.
- $\mathcal{B}(S)$, $\mathcal{B}(E)$ the Borel $\sigma-$algebra of S or E, respectively.
- $\mathcal{W}_T^d$ denotes the normed space continuous functions $C([0, T]; \mathbb{R}^d)$ with the uniform norm.
- $\mathcal{W}^d$ denotes the metric space of continuous functions $C([0, \infty); \mathbb{R}^d)$ with a metric inducing the uniform convergence over compact sets.
- $V_T^{x,\phi}$ denotes the portfolio value process with initial cost x and a trading strategy ϕ.
- Ξ_t is the set of $\mathcal{F}_t-$measurable random variables.
- $\mathscr{W}$ the familiy of measurable functions with finite moments of all order.
- $D_t(\omega)$ the smallest affine subspace (or linear manifold) containing the support of the regular conditional distribution of ΔS_t given $\mathcal{F}_{t-1}$.
- $\mathcal{M}^e(S)$ denotes the set of equivalent martingale measures for S.
- S_+^d denotes the set of positive semidefinite symmetric matrices.
- $\mathcal{A}^{d,r}$ denotes the set of adapted and jointly measurable functionals on $[0, \infty) \times \mathcal{W}^d$. See Definition 4.2.7 for details.

- $C_b^2\left(\mathbb{R}^d;\mathbb{R}\right)$ the space of bounded continuous functions with continuous second order derivatives.
- $C_0^\infty\left(\mathbb{R}^d\right)$ the space of smooth functions with compact support.
- $C_b^{1,2}\left(\mathbb{R}_+\times\mathbb{R}^d;\mathbb{R}\right)$ is the space of continuous differentiable functions of order 1 in time and order 2 in space.
- α_t the truncation operator in $\mathcal{W}^d$.
- $\mathbb{A}=S_+^d\times\mathbb{R}^d$.
- $A_t\left(x.\right)$, $A_t\left(\omega\right)$ a class of convex subsets of $\mathbb{A}$.
- $tr\left(a\right)$ denotes the trace of a matrix a.
- $F_t\left(u,v,\omega\right)$, $F_t\left(u,v,x.\right)$ denotes the support functions of the convex sets $A_t\left(\omega\right), A_t\left(x.\right)$, respectively.
- $\operatorname{ess\,inf}_{i\in I} X_i$ and $\operatorname{ess\,sup}_{i\in I} X_i$ the essential infimum and essential supremum of a collection of random variables $\left\{X_i\right\}_{i\in I}$.

Contents

Chapter 1

INTRODUCTION

The main subject of this work is the analysis of a broad class of portfolio optimisation problems in incomplete markets under the Cumulative Prospect Theory (CPT) developed by Kahneman D. and Tversky A. in [30] and [57]. We study a well-known mathematical formulation of the problem of optimal investment and answer some questions concerning the existence of optimal strategies and well-posedness. The framework is robust enough to encompass the problems of Expected Utility Theory and the generalisations related to utility maximisation when investors are loss averse and risk seeking.
Nevertheless, the generality of our setting comes at the expense of its mathematical tractability. First of all, even though choices can be assessed by means of an objective function, it is unclear that, in general, the problem of optimising such a functional is a well-defined optimisation problem, we give some conditions that ensure that the problem is well-posed. Secondly, the traditional methods, such as dynamic programing and HJB equations and duality theory are no longer applicable.

While at first glance, the description and analysis of this problem may seem simply a zeal of mathematical generalisation, this is not the case. The axioms on which the Expected Utility Theory (EUT) is based have been questioned systematically. One of the ideas underlying this theory is the assumption that investors are *risk-averse*, this means, that as their wealth grows, its increase in 'satisfaction' on this favorable change becomes smaller. In other words, the utility function to evaluate total wealth is a concave function. On the other hand, EUT axioms imply that investors are fully rational and hence, all economic agents have a 'transparent' and objective view of the outcomes (or rather, the likelihood of the outcomes). Kahneman and Tversky's findings suggested that such an assumption is not consistent with people's approach to decision making. In a series of studies ([58], [30] and [57]) they show that investors tend to magnify small probabilities and underestimate big probabilities. In other words, investors have a distorted perception of the distributions of their wealth. Therefore under the tenets of the CPT, an investor is no longer rational.
Another element proposed by Kahneman and Tversky, that should be pointed out is the claim that economic agents evaluate their investments according to their own levels of wealth rather than absolute levels of wealth. Thus, every investor has a reference point that allows to define their relative 'losses' and 'gains'. This element explains why, given an equal absolute value of wealth, two investors can take different decisions and have different views of the market.

A brief outline of how this work is organised follows. In chapter 2, important results on weak convergence of probability measures; martingale theory; and discrete-time finance models are described; we investigate a result proved in [44] and a well-known theorem in finance literature from [27]. These facts form the main background to introduce in chapter 3 the problem of optimal investment under the CPT theorem in a discrete-time setting, this is based upon our publication [42] and [11]. We describe our model, present some assumptions and main results are derived.

The second part of this work comprises the problem of optimal investment in continuous-time setting, we follow closely [25] and [35], we descibe the martingale problem formulation of diffusion processes in Chapter 4. A key result on the limits and topological properties of the set of laws of a class of Itô processes is described in Chapter 5, the aim of this chapter is to give an explanation of the results in [32]. Finally in Chapter 6, we describe a factor model that includes a class of stochastic volatility models possibly with path-depending coefficients, under this model, the problem of optimal investment with a behavioural investor is described, results on well-posedness and existence of optimal strategies under the framework of weak solutions are provided. This chapter is based upon our second contribution, [43].

Further research and possible limitations of our techniques motivate the conclusion. The work concludes with some of the future research problems.

Chapter 2

PRELIMINARIES

1 Some auxiliary results on weak convergence.

In this section we introduce some notions of convergence of probability measures, we follow closely the standard references [6] and [22].
Throughout this chapter, we assume that S is a Polish metric space, hereafter $\mathcal{B}(S)$ is the Borel σ-algebra on S.
We denote by $\mathcal{M}^1(S)$ the set of probability measures on $\mathcal{B}(S)$, i.e. the set of probability measures on the Borel $\sigma-$algebra on S. It is possible to define a metrizable topology τ_w on $\mathcal{M}^1(S)$ such that $\left(\mathcal{M}^1(S), \tau_w\right)$ is a Polish space, for proofs of these facts see [51], Theorem 9.15.

Definition 2.1.1. We say that a sequence of probability measures $\{\mu_n\}_{n\geqslant 1}$ converges weakly to a probability measure μ if for any real-valued, bounded continuous function f we have

$$\lim_{n\to\infty} \int_S f(x)\, \mu_n(dx) = \int_S f(x)\, \mu(dx). \tag{2.1.1}$$

We indicate this convergence by writing $\mu_n \Rightarrow \mu$.

Condition in (2.1.1) enables to formulate the weak topology through a family of 'basic' open sets, let $f_i \in C_b(S;\mathbb{R})$ and $\mu \in \mathcal{M}^1(S)$

$$U_r(\mu; f_1, \dots f_n) := \left\{\nu \in \mathcal{M}^1(S) \,:\, \left|\int_S f_i(s)\,\mu(ds) - \int_S f_i(s)\,\nu(ds)\right| < r,\ i = 1, \dots, n\right\}$$

Let $\mathcal{U} := \left\{U_r(\mu; f_1, \dots, f_r),\, r > 0, \{f_i\}_{i\geqslant 1}^n, n \in \mathbb{N}\right\}$. A base at the point $\mu \in \mathcal{M}^1(S)$ of the weak topology is then defined to be all the sets that are union of finite intersections of sets $U_r \in \mathcal{U}$.
The following concept is central in the theory of weak convergence of probability measures.

Definition 2.1.2. A family of probability measures $\{\mathbb{P}\}_{\alpha\in I} \subset \mathcal{M}^1(S)$ is tight (or uniformly tight) if for any $\epsilon > 0$ there exists a compact set K_ϵ such that

$$\inf_{\alpha\in I} \mathbb{P}_\alpha(K_\epsilon) \geqslant 1 - \epsilon.$$

Definition 2.1.3. A subclass $\mathcal{A} \subset \mathcal{B}(S)$ is a *separating class* if two probability measures that agree on $\mathcal{A}$ necessarily agree on $\mathcal{B}(S)$.

Example 2.1.4. On $(S, \mathcal{B}(S))$, the family of closed sets is a separating class due to the $\pi-$system Lemma (or Dynkin's Lemma) and the fact that the family of closed sets generates the Borel σ-algebra.

Example 2.1.5. An important example of a Polish space is the space of continuous functions on a finite interval, $\mathcal{W}_T^d := C\left([0,T];\mathbb{R}^d\right)$ with the uniform convergence norm $\|x\| = \sup_{t\in[0,T]}|x(t)|$. We denote by $\mathcal{B}\left(\mathcal{W}_T^d\right)$ its Borel σ-algebra.
Suppose $\{t_1, t_2, \dots, t_k\} \subset [0,T]$ is a collection of points. On $\mathcal{W}_T^d$, we define the projections $p_{t_1,t_2,\dots,t_k} : \mathcal{W}_T^d \to \mathbb{R}^{dk}$ as the functions that put into correspondence any continuous function x with their values on the points $t_1, t_2, \dots, t_k$ i.e. $p_{t_1,t_2,\dots,t_k}(x) := (x(t_1), x(t_2), \dots, x(t_k))$.

Definition 2.1.6. In $\mathcal{W}_T^d$, we define the *finite-dimensional rectangles* or tubes to be the sets of the form $p^{-1}_{(t_1,t_2,\dots,t_k)}(H)$ for $H \in \mathcal{B}\left(\mathbb{R}^{dk}\right)$. Such a class of rectangles is denoted by $\mathcal{C}_f$.

Clearly, the family of projections are continuous mappings on $\mathcal{W}_T^d$, hence, each finite-dimensional rectangle is Borel measurable. This class is clearly a $\pi-$system [1] and any ball in $\mathcal{W}_T^d$ is a countable intersection of elements of $\mathcal{C}_f$.

$$\overline{B(x;\epsilon)} = \bigcap_{r\in\mathbb{Q}\cap[0,T]} \{y : |y(r) - x(r)| \leqslant \epsilon\},$$

thus, the $\sigma-$field $\sigma(\mathcal{C}_f)$ contains the closed balls, hence it contains the Borel sets in $\mathcal{W}_T^d$. As $\mathcal{C}_f$ is a π-system, then the finite-dimensional tubes is a separating class.

The following important fact characterising weak convergence is widely used, we include it for the sake of completeness. It is referred as the 'portmanteau theorem'.

Proposition 2.1.7. *Let $(S, \mathcal{B}(S))$ be a separable metric space and let $\{\mu_n\}_{n\geqslant 1} \subset \mathcal{M}^1(S)$ be a sequence of probability measures, on $\mathcal{B}(S)$. Then the following statements are equivalent*

[I] A sequence of probability measures $\{\mu_n\}_{n\geqslant1}$ converges weakly to a measure μ.

[II] For any closed set F
$$\overline{\lim_{n\to\infty}}\mu_n(F) \leqslant \mu(F). \tag{2.1.2}$$

[III] For any open set U
$$\mu(U) \leqslant \underline{\lim_{n\to\infty}}\,\mu_n(U). \tag{2.1.3}$$

[IV] If G is a Borel set such that its boundary ∂G has zero μ-measure then
$$\lim_{n\to\infty}\mu_n(G) = \mu(G). \tag{2.1.4}$$

[V] For any bounded function f such that the set of discontinuity points D_f has $\mu-$measure zero then
$$\lim_{n\to\infty}\int_S f(s)\,\mu_n(ds) = \int_S f(s)\,\mu(ds). \tag{2.1.5}$$

Proof. $[I] \implies [II]$. Suppose $\mu_n \Rightarrow \mu$, let F be a closed set of S and define the following sequence of continuous functions $f_m(x) := \frac{1}{1+md(x,F)}$, the function is equal to 1 if $x \in F$ and

[1] A $\pi-$system say Π is a family of subsets, of S that it is stable under finite intersection of sets, i.e. if $F, H \in \Pi$ then $F \cap H \in \Pi$

$0 \leqslant f \leqslant 1$ for all $x \in S$, moreover if $x \notin F$ we have that $f_m(x) \downarrow 0$ hence $f_m(x) \downarrow \mathbb{I}_F(x)$. By assumption

$$\overline{\lim_{n\to\infty}}\mu_n(F) \leqslant \overline{\lim_{n}} \int_S f_m(x)\,\mu_n(dx) = \int_S f_m(x)\,\mu(dx)\,, \tag{2.1.6}$$

as the left hand side of the inequality in (2.1.6) does not depend on m, by the bounded convergence theorem, letting $m \to \infty$; we obtain the required conclusion.

$[II] \iff [III]$. Suppose that (2.1.2) holds. Let U be an open subset of S, then applying *[II]* to S/U

$$\overline{\lim_{n\to\infty}}\big(1-\mu_n(U)\big) = \overline{\lim_{n\to\infty}}\mu_n(S/U) \leqslant \mu(S/U) = 1-\mu(U)\,,$$

simplifying the last expression yields $\underline{\lim}_n \mu_n(U) \geqslant \mu(U)$. Similarly, one can see that (2.1.3) implies (2.1.2) thus *[II]* and *[III]* are equivalent.

$[I] \Longrightarrow [IV]$. As explained before, *[I]* implies *[II]* and *[III]*. Notice that $\overline{G}/int(G) = \partial G$ then by *[II]* and *[III]* then

$$\overline{\lim_{n\to\infty}}\,\mu_n\left(\overline{G}\right) \leqslant \mu\left(\overline{G}\right) = \mu\left(int(G)\right) \leqslant \underline{\lim_{n\to\infty}}\,\mu_n\left(int(G)\right),$$

as we have also that $\underline{\lim}_n\,\mu_n(int(G)) \leqslant \underline{\lim}_n\,\mu_n(G)$ and $\overline{\lim}_n\,\mu_n(G) \leqslant \overline{\lim}_n\,\mu_n\left(\overline{G}\right)$ the claim follows.

$[IV] \Longrightarrow [V]$. Let f be a function with $\mu(D_f) = 0$, denote by $M := \|f\|_\infty$ and $A_t = \{s \in S : f(s) \geqslant t\}$ then $\partial A_t \subseteq D_f \cup \{f(s) = t\}$. On the other hand $F(t) = \mu(\{f \geqslant t\})$ is decreasing and left continuous, then there is at most a countable set of points $\{t_i\}_{i\geqslant 1}$ such that $\mu(\{s \in S :\, f(s) = t_i\}) > 0$. This implies

$$\int_0^\infty \mu\left(\overline{A_t}\right)dt = \int_0^\infty \mu\left(intA_t \cup \partial A_t\right)dt = \int_0^\infty \mu(A_t)\,dt, \tag{2.1.7}$$

$$\int_S f(s)\,\mu(ds) = \int_S\int_0^\infty \mathbb{I}_{\{f(s)\geqslant t\}}dt\mu(ds) = \int_0^\infty \mu(A_t)\,dt,$$

As f is bounded, for $t > M$, $\mu(\{s :\, f(s) > t\}) = 0$, as μ is a finite measure we can apply the 'reverse' Fatou's lemma, and (2.1.7)

$$\begin{aligned}\overline{\lim_{n\to\infty}}\int_0^\infty \mu_n(A_t)\,dt \leqslant \int_0^\infty \overline{\lim_{n\to\infty}}\,\mu_n(A_t)\,dt \leqslant \int_0^\infty \mu\left(\overline{A}_t\right)dt = \\ \int_0^\infty \underline{\lim_{n\to\infty}}\,\mu_n(A_t)\,dt \leqslant \underline{\lim_{n\to\infty}}\int_0^\infty \mu_n(A_t)\,dt,\end{aligned} \tag{2.1.8}$$

and the claim follows as

$$\overline{\lim_{n\to\infty}}\int_0^\infty \mu_n(A_t)\,dt = \overline{\lim_{n\to\infty}}\int_S f(s)\,\mu_n(ds) \;\text{ and }\; \underline{\lim_{n\to\infty}}\int_0^\infty \mu_n(A_t)\,dt = \underline{\lim_{n\to\infty}}\int_S f(s)\,\mu_n(ds)\,. \tag{2.1.9}$$

$[V] \Longrightarrow [I]$ If $f \in C_b(S;\mathbb{R})$ then f satisfies the condition $(\mu(D_f) = 0)$ thus,

$$\lim_{n\to\infty}\int_S f(s)\,\mu_n(ds) = \int_S f(s)\,\mu(ds)\,.$$

$\square$

Notation Suppose that X and Y are Polish spaces and $h : X \to Y$ is a Borel measurable map (i.e. $\mathcal{B}(X)/\mathcal{B}(Y)$-measurable), recall that any measure Q on $\mathcal{B}(X)$ induces a measure on

$\mathcal{B}(Y)$ through the map h. Let $A \in \mathcal{B}(Y)$ denote by $Q \circ h^{-1}(A) := Q\left(h^{-1}(A)\right)$.
The convergence in the weak topology in $\mathcal{M}^1(S)$ defines another type of convergence of random variables. We follow closely [6].

Definition 2.1.8. Let $\{X_n\}_{n \geqslant 1}$ be a sequence of $\mathbb{R}^d-$valued random variables, we say that the sequence $\{X_n\}_{n \geqslant 1}$ converges in distribution to the random variable X if their laws $\mu_n := \mathbb{P} \circ X_n^{-1}$ converge weakly to $\mu := \mathbb{P} \circ X^{-1}$. In our notation, $\mathbb{P} \circ X_n^{-1} \Rightarrow \mathbb{P} \circ X^{-1}$.

If $S = \mathbb{R}$, Definition 2.1.8 can be rephrased in terms of the convergence of distribution functions.

Proposition 2.1.9. *Suppose that* $\{\mu\}, \{\mu_n\}_{n \geqslant 1} \subset \mathcal{M}^1(\mathbb{R})$, *are the laws of random variables* $\{X\}, \{X_n\}_{n \geqslant 1}$, *respectively; let* $F(x)$ *and* $\{F_n(x)\}_{n \geqslant 1}$ *be their associated distribution functions. The sequence of laws of* X_n *converges to the law of* X *i.e.* $\mu_n \Rightarrow \mu$ *if and only if* $F_n(x) \to F(x)$ *for all continuity points of* F.

The reader is referred to [6] and [22] for proofs of this fact.

Theorem 2.1.10. *Suppose that* h *maps* S *into* $\mathbb{R}^n$. *Further, suppose that* $\mathbb{P}_n \Rightarrow \mathbb{P}$. *If* h *is a continuous function then* $\mathbb{P}_n \circ h^{-1} \Rightarrow \mathbb{P} \circ h^{-1}$.
Furthermore, suppose h *is a function such that the set of discontinuity points of* h, D_h, *is such that* $\mathbb{P}(D_h) = 0$ *then* $\mathbb{P}_n \circ h^{-1} \Rightarrow \mathbb{P} \circ h^{-1}$.

Proof. Let F be a closed set in $\mathbb{R}^n$. By assumption $\mathbb{P}(D_h^c) = 1$. Then

$$\limsup_n \mathbb{P}_n\left(h^{-1}(F)\right) \leqslant \limsup_n \mathbb{P}_n\left(\overline{h^{-1}(F)}\right) \leqslant \mathbb{P}\left(\overline{h^{-1}(F)}\right) = \mathbb{P}\left(D_h^c \cap \overline{h^{-1}(F)}\right) \leqslant \mathbb{P}\left(h^{-1}(F)\right). \tag{2.1.10}$$

the second inequality is a consequence of portmanteau theorem, Proposition 2.1.7, the equality is a consequence of the theorem's assumption. The last inequality is a consequence of the inclusion $D_h^c \cap \overline{h^{-1}(F)} \subset h^{-1}\left(\overline{F}\right)$ which we shall prove, let $x \in \overline{h^{-1}(F)}$ then there is a sequence $\{x_n\}_{n \geqslant 1} \subset S$ with $h(x_n) \in F$. Further if $x \in D_h^c$ as well, then $h(x_n) \to h(x)$ and $h(x) \in \overline{F}$ thus, $x \in h^{-1}\left(\overline{F}\right)$. In other words, $\overline{h^{-1}(F)} \cap D_h^c \subset h^{-1}\left(\overline{F}\right)$. Then (2.1.10) yields in terms of the induced probability measures the following

$$\limsup_n \mathbb{P}_n \circ h^{-1}(F) \leqslant \mathbb{P} \circ h^{-1}(F),$$

by portmanteau theorem this implies that $\mathbb{P}_n \circ h^{-1} \Rightarrow \mathbb{P} \circ h^{-1}$. □

2 Prokhorov's theorem and related results

Prokhorov's theorem is an important tool used throughout probability theory and its applications. This theorem replaces in some sense the well-known 'Bolzano-Weierstrass' theorem (valid in finite-dimensional spaces) when dealing with probability measures. The boundedness condition (in the case of Bolzano-Weierstrass) is replaced by a condition of uniform tightness of the family of probability measures, see Definition 2.1.2. This theorem is extensively used to prove limit theorems. The main reference for applications of the theory of weak convergence to the theory of stochastic process is [26].

Definition 2.2.1. We say that a family of probability measures Γ on a measurable space $(S, \mathcal{B}(S))$, is *weakly relatively compact* if every sequence of probability measures $\{\mathbb{P}_n\}_{n\geqslant 1} \subset \Gamma$ has a weakly convergent subsequence.

The concept of relatively compactness is a useful tool to prove claims on the law of processes, more specifically suppose that $X_n : \Omega \to \mathcal{W}_T^d$ is a sequence of processes with values in $\mathbb{R}^d$ and $X : \Omega \to \mathcal{W}_T^d$, in general proving that $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$ is a difficult task, a possible method is given by the following remark.

Remark 2.2.2. Suppose that on $\mathcal{B}\left(\mathcal{W}_T^d\right)$ there is a sequence of probability measures $\mathbb{P}_n$ and $\mathbb{P}$ on this measurable space such that for all k, $\{t_1, t_2, \ldots, t_k\} \subset [0, T]$

$$\mathbb{P}_n \circ p_{t_1,t_2,\ldots,t_k}^{-1} \Rightarrow \mathbb{P} \circ p_{t_1,t_2,\ldots,t_k}^{-1}. \tag{2.2.1}$$

If, in addition, the sequence of probability measures is relatively compact by Theorem 2.1.10 then there is a subsequence of probabiliy measures $\{\mathbb{P}_{n_k}\}$ such that $\mathbb{P}_{n_k} \circ p_{t_1,t_2,\ldots,t_k}^{-1} \Rightarrow \mathbb{Q} \circ p_{t_1,t_2,\ldots,t_k}^{-1}$ as weak limits are unique it follows that $\mathbb{Q} \circ p_{t_1,t_2,\ldots,t_k}^{-1} = \mathbb{P} \circ p_{t_1,t_2,\ldots,t_k}^{-1}$ for any k integer and any collection of points $\{t_1, t_2, \ldots, t_k\} \subset [0, T]$. Thus, $\mathbb{P}$ and $\mathbb{Q}$ are equal on the class of finite dimensional rectangles $\mathcal{C}_f$ (see Definition 2.1.3) then $\mathbb{P} = \mathbb{Q}$. Then each subsequence has a further subsequence that converges weakly to $\mathbb{P}$ hence $\mathbb{P}_n \Rightarrow \mathbb{P}$.

In other words, convergence of finite-dimensional distributions and weakly relative compactness of the set of laws allow to show that the laws of the processes converge weakly. The following result is known as Prokhorov's theorem, please refer to [39].

Theorem 2.2.3. *Let $\Gamma \subset \mathcal{M}^1(S)$ be a family of probability measures on S a Polish space then Γ is relatively compact with respect the weak topology if and only if Γ is tight.*

For a proof of this important theorem, the reader is referred to [6] Theorem 1.5.1 and Theorem 1.5.2 or the article [39].

Example 2.2.4. Tightness of the family of probability measures is a required condition to have relatively compactness. For instance, on $\mathcal{B}(\mathbb{R})$ consider the measures given by $\nu_n(\cdot) := \delta_n(\cdot)$ is a sequence of probability measures that has no convergent subsequence.

Indeed, if there was a converging subsequence $\nu_{n_k} \Rightarrow \eta$, let us denote by $F_\eta(x)$ its distribution function, for fixed $x \in \mathbb{R}^+$ by Proposition 2.1.7 *III* we would have

$$F_\eta(x_-) = \eta((-\infty, x)) \leqslant \underline{\lim}_n \nu_n((-\infty, x)) = 0.$$

And outside of a countable set, denoted by D, $F_\eta(x) = 0$. Since the set of discontinuity points is countable, its complement is dense, if x_n is a discontinuity point then it is possible to take a sequence $\{y_n^k\} \subset \mathbb{R}^+/D$ such that $y_n^k \downarrow x_n$ then we have $F_\eta(x_n) = 0$ but this contradicts the fact that F_η was a distribution function, and hence ν_n does not have a convergent sequence. One sees that the total mass 'evaporates' as $n \to \infty$. Tightness in Theorem 2.2.3 prevents this phenomenon. Some standard properties of tight distributions are listed below.

Lemma 2.2.5. *Let $\left(X^k\right)_{k\geqslant 1}$ be sequence of random variables (i.e. the laws of X^k are tight) in $\mathbb{R}^n$, let Y be a $\mathbb{R}^n$-valued random variable. If the sequence of laws of $\left(X^k\right)_{k\geqslant 1}$ is tight then*

i) *The sequence* $\left\{\left(X^k, Y\right)\right\}_{k\geqslant 1}$ *is a tight sequence of random variables in* $\mathbb{R}^{2n}$.

ii) *The sequence* $\left\{X^k + Y\right\}_{k\geqslant 1}$ *is tight.*

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $X, X^1,\ X^2 \ldots,$ are real-valued random variables on this probability space. Recall that we say that the sequence $\{X^n\}_{n\geqslant 1}$ converges *in probability* to X if for any $\eta > 0$ and any $\epsilon > 0$ there is a N such that for any $n \geqslant N$

$$\mathbb{P}\left(|X^n - X| > \eta\right) < \epsilon,$$

or briefly $\lim_{n\to\infty} \mathbb{P}\left(|X^n - X| > \eta\right) = 0$ for any $\eta > 0$.
The following theorem will be applied and it is a useful fact on convergence in probability. It is stated in [3].

Theorem 2.2.6. *Let* (S, d) *be a separable metric space and let* h *be a Borel function from* S *to* $\mathbb{R}^d$. *Let* $\left\{X^k\right\}_{k\in\mathbb{N}}$ *be a sequence of random variables with values in* (S, d) *and suppose that such a sequence converges in probability to* X, *suppose in addition that the random variables* X^k *have the same law (i.e. the sequence* $\left\{X^k\right\}_{k\in\mathbb{N}}$ *are identically distributed) then* $h \circ X^k$ *converges to* $h \circ X$ *in probability.*

Proof of Theorem 2.2.6. In the case h is a continuous function, the claim follows from a well-known condition for a sequence to converge in probability. To wit, a sequence of random variables $\left\{X^k\right\}_{k\geqslant 0}$ converges in probability if and only if for any subsequence there is a further subsequence $\left\{X^{k_m}\right\}_{m\geqslant 0}$ such that X^{k_m} converges to X $\mathbb{P}-a.s.$ In this case, $h \circ X^{k_m}$ converges to $h \circ X$ $\mathbb{P} - a.s.$
For the general case, notice that the measure $\mu : \mathcal{B}(S) \to [0, 1]$, defined by $\mu(A) = \mathbb{P}(X \in A)$ induces a probability measure on $\mathcal{B}(S)$ and furthermore, this measure is closed regular[2]. By Luzin's theorem, Theorem 7.5.2 in [22] there is a closed set F_n such that

- The function h restricted to F_n is continuous.
- For each n we have $\mu(S/F_n) < \frac{1}{2^{n+1}}$.

Denote by $D_n := \bigcap_{k=n}^{\infty} F_k$, the sets are closed and $\mu(S/D_n) \leqslant \sum_{k=n}^{\infty} \mu(S/F_n) \leqslant \frac{1}{2^n}$. The mapping h restricted to D_n is continuous. Applying Tietze's extension theorem to h, Theorem 2.6.4 in [22], for each $n = 1, \ldots$ we can find a continuous function $h_n : S \to \mathbb{R}^d$ such that $h\Big|_{F_n} = h_n\Big|_{F_n}$. And it follows that for any $\eta > 0$

$$\mu\left\{|h(x) - h_n(x)| > \eta\right\} \leqslant \mu(S/D_n) \leqslant \frac{1}{2^n}. \tag{2.2.2}$$

The Borel-Cantelli lemma implies that $h_n(x) \to h(x)$ $\mu - a.s.$ This implies our claim, indeed, taking such functions h_n defined above

$$\mathbb{P}\left(\left|h\left(X^k\right) - h(X)\right| > 3\eta\right) \leqslant \mathbb{P}\left(\left|h\left(X^k\right) - h_n\left(X^k\right)\right| > \eta\right) + \mathbb{P}\left(\left|h_n\left(X^k\right) - h_n(X)\right| > \eta\right) + \tag{2.2.3}$$
$$\mathbb{P}\left(|h_n(X) - h(X)| > \eta\right).$$

[2]A measure μ is *closed regular* if for any Borel set we have $\mu(A) = \sup\{\mu(F) : F \subset A, \text{ and } F \text{ is closed}\}$

Let $\delta > 0$, by our assumptions, the last sum of (2.2.3) is equal to the first term and then

$$\mathbb{P}\left(\left|h\left(X^k\right) - h\left(X\right)\right| > 3\eta\right) \leqslant 2\mathbb{P}\left(\left|h\left(X\right) - h_n\left(X\right)\right| > \eta\right) + \mathbb{P}\left(\left|h_n\left(X^k\right) - h_n\left(X\right)\right| > \eta\right), \quad (2.2.4)$$

As proved before, and since a.s. convergence implies convergence in probability, for any $\delta > 0$ there exists N such that for $n \geqslant N$, $\mathbb{P}\left(\left|h\left(X\right) - h_n\left(X\right)\right| > \eta\right) < \delta$. Fix $n_0 \geqslant N$ as h_{n_0} is continuous, $h_{n_0}\left(X^k\right) \to h_{n_0}\left(X\right)$ in probability as $k \to \infty$. Therefore, choosing k large enough we have

$$\mathbb{P}\left(\left|h\left(X^k\right) - h\left(X\right)\right| > 3\eta\right) \leqslant 2\mathbb{P}\left(\left|h\left(X\right) - h_n\left(X\right)\right| > \eta\right) + \mathbb{P}\left(\left|h_n\left(X^k\right) - h_n\left(X\right)\right| > \eta\right) \leqslant \delta,$$

and the claim follows. □

For a generalisation of Theorem 2.2.6 see Theorem 1, p.314 in [17].
Finally, we review some results on the generation of independent and uniformly distributed random variables. These lemmas are needed in chapter 3, the reader is referred to [11] for proofs of these facts.

Lemma 2.2.7. *Let ε be a uniformly distributed on $[0, 1]$. Then for each $k \geqslant 1$ there are Borel measurable functions f_1, f_2, ..., $f_l : [0, 1] \to [0, 1]$ such that $f_1(\varepsilon), \ldots f_k(\varepsilon)$ are independent and uniformly distributed on $[0, 1]$.*

Lemma 2.2.8. *Let $\mu(dz, dy)$ be a probability measure on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that we can 'disintegrate' $\mu(dy, dz) = \nu(y, dz)\,\delta(dz)$ and $\delta(dz)$ is a probability measure on $\mathbb{R}^{n_2}$ and $\nu(y, dz)$ is a kernel in the sense of Definition A.1.1. Assume that Y has distribution $\delta(dy)$ and U is a r.v. independent of Y and uniformly distributed on $[0, 1]$. Then there is a measurable function $G : \mathbb{R}^{n_2} \times [0, 1] \to \mathbb{R}^{n_1}$ such that $(Y, G(Y, U))$ its distribution is equal to μ.*

Lemma 2.2.9. *Let (X, U) be and $n+m-$dimensional random variable such that the conditional law of X given U is a.s. atomless. Then there is a measurable $G : \mathbb{R}^{n+m} \to \mathbb{R}^n$ such that $G(X, U)$ is independent of U with uniform law on $[0, 1]$.*

3 Relevant results on martingale theory

We recall some definitions that although, they belong to the theory of stochastic processes, such concepts have a direct relevance into modern mathematical finance.
We present for the sake of completeness some results that are used in the following chapter. In a discrete-time setting, some families of processes involving martingales and local martingales happen to be the same. We shall consider a fixed filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathbb{P}\right)$ and $\mathcal{F}$ is $\mathbb{P}-$complete. Recall the following definitions.

Definition 2.3.1. We say that a stochastic process $\{X_n, \mathcal{F}_n\}_{n\geqslant 0}$ is a *local martingale* if it is an $\mathcal{F}_n-$adapted process and if there is an increasing sequence of $\mathcal{F}_n-$stopping times $\{\tau_k\}_{k\geqslant 1}$ such that $\mathbb{P}(\tau_k \nearrow \infty) = 1$ and $X^{\tau_k} = \{X_{\tau_k \wedge n}, \mathcal{F}_n\}$ is a martingale for any k.[3].

[3] In the case when the time index is a finite set, the condition on the stopping times is replaced by $\mathbb{P}(\tau_k \geqslant N) \to 1$

Definition 2.3.2. We say that a stochastic process $\{X_n, \mathcal{F}_n\}_{n\geqslant 0}$ is a *generalised martingale* if it is an $\mathcal{F}_n$-adapted process, $\mathbf{E}\left[|X_{n+1}|\,|\mathcal{F}_n\right] < \infty$ $\mathbb{P}-$a.s. and

$$\mathbf{E}\left[X_{n+1}\,|\mathcal{F}_n\right] = X_n,\ \mathbb{P}-\text{a.s.} \tag{2.3.1}$$

The conditional expectation in (2.3.1) is defined on $\{\mathbf{E}\left[|X_{n+1}|\,|\mathcal{F}_n\right] < \infty\}$ as $\mathbf{E}\left[X_{n+1}\,|\mathcal{F}_n\right] = \mathbf{E}\left[X^+_{n+1}\,|\mathcal{F}_n\right] - \mathbf{E}\left[X^-_{n+1}\,|\mathcal{F}_n\right]$.

Definition 2.3.3. Suppose $\{M_n, \mathcal{F}_n\}_{n\geqslant 0}$ is a $d-$dimensional martingale and $\{\gamma_n\}_{n\geqslant 1}$ is a $d-$dimensional, finite valued and $\mathcal{F}_n-$predictable processes. We say that X_n is a *d-martingale transform* if the process X_n is of the form

$$X_k = X_0 + \sum_{i=1}^{k} \langle \gamma_i, \Delta M_i \rangle . \tag{2.3.2}$$

Definition 2.3.4. We say that a process $\{X_n, \mathcal{F}_n\}_{n\geqslant 0}$ is a *$d-$local martingale transform* if the process $\{M_n\}$ in (2.3.2) is a $d-$dimensional local martingale.

Definition 2.3.5. We denote by $\mathcal{M}_{loc}$, GM, MT^d and LMT^d the classes of local martingales, generalised martingales, $d-$dimensional martingale transform and $d-$dimensional local martingale transform respectively.

It turns out that in a discrete-time setting and under some conditions, the notions of local martingales, d-martingale transform and d-local martingale transforms coincide.

Theorem 2.3.6. *Suppose $\left(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathbb{P}\right)$ is a filtered probability space, then the classes $\mathcal{M}_{loc}$, MT^d and LMT^d coincide. Moreover $\mathcal{M}_{loc} \subset GM$.*

Proof. $LMT^d \subset \mathcal{M}_{loc}$. Suppose $X \in LMT^d$, by definition, there is a sequence of $\mathcal{F}_n-$stopping times $\{\sigma_n\}_{n\geqslant 1}$ such that M^{σ_n} are martingales. Define $\tau_n := \sigma_n \wedge \inf\{t \geqslant 0 : |\gamma_t| > n\}$, as γ is predictable $\{\inf\{t \geqslant 0 : |\gamma_t| > n\} = N\} = \bigcap_{i=1}^{N-1} \{|\gamma_i| < n\} \cap \{|\gamma_N| \geqslant n\} \in \mathcal{F}_N$, thus, τ_k is a stopping time for any $k \geqslant 0$ and $a.s.$ increasing to ∞.
From the usual identity

$$X^{\tau_k}_n = X_{\tau_k\wedge n} = X_0 + \sum_{j=1}^{\tau_k\wedge n} \langle \gamma_j, (M_{j+1} - M_j) \rangle = X_0 + \sum_{j=1}^{n} \left\langle \gamma_j \mathbb{I}_{(j\leqslant\tau_k)}, (M_{j+1} - M_j) \right\rangle,$$

and

$$X^{\tau_k}_n = X_0 + \sum_{j=1}^{n} \left\langle \mathbb{I}_{(j\leqslant\tau_k)}\gamma_j, \left(M^{\tau_k}_{j+1} - M^{\tau_k}_j\right) \right\rangle,$$

and as $M^{\tau_k}_{n+1} - M^{\tau_k}_n$ is a martingale (and $\tau_k \leqslant \sigma_k$), the sum in the last equation is also a martingale since $\gamma'_i = \mathbf{I}_{\{i\leqslant\tau_k\}}\gamma_i$ is bounded and predictable. Then $X^{\tau_n} \in MT^d$ (in particular a martingale) and X is a local martingale.
$M_{loc} \subset GM$. Suppose $X \in \mathcal{M}_{loc}$, then there is a null set $N \in \mathcal{F}_0$ such that, $\tau_k \nearrow \infty$ on Ω/N and X^{τ_k} is a martingale. Thus for any n, $\Omega/N = \bigcup_{k=1}^{\infty} \{\tau_k > n\}$. On the set $\{\tau_k > n\}$ we have $\mathbf{E}\left[\left|X^{\tau_k}_{n+1}\right|\,|\mathcal{F}_n\right] = \mathbf{E}\left[|X_{n+1}|\,|\mathcal{F}_n\right] < \infty$ and $\mathbf{E}\left[X^{\tau_k}_{n+1}\,|\mathcal{F}_n\right] = \mathbf{E}\left[X_{n+1}\,|\mathcal{F}_n\right] = X_n = X^{\tau_k}_n$ hold. Thus, $X \in GM$.
Clearly, $MT^1 \subset MT^d \subset LMT^d$.

Every local martingale $\{X_n\}_{n\geqslant 1}$ is also a local martingale transform (by taking the predictable and constant process $\gamma_i := 1$ for all $1 \leqslant i \leqslant N$). We conclude that the families of processes are the same. □

We follow the original source, [27]. If $\mathcal{F}_0 = \{\varnothing, \Omega\}$ we have a stronger result.

Theorem 2.3.7. *Suppose* $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 1}, \mathbb{P}\right)$ *is a filtered probability space and* $\mathcal{F}_0 = \{\varnothing, \Omega\}$ *then the classes* $\mathcal{M}_{loc}$*,* GM*,* MT^d *and* LMT^d *coincide.*

Proof. • $GM \subset LMT^d$. And this is enough to show that such classes of processes are the same. Let $\{X_n, \mathcal{F}_n\}_{n\geqslant 1}$ and suppose $X \in GM$. Define the sets

$$A(n,k) := \{\omega \, : \, \mathbf{E}\left[|\Delta X_{n+1}| \,|\mathcal{F}_n\right] \in [k, k+1)\}, \tag{2.3.3}$$

and define $u_n := \sum_{k\geqslant 0} \frac{1}{(k+1)^3} \Delta X_n \mathbb{I}_{A(n-1,k)}$, the process is such that $\mathbf{E}\left[|u_n| \,|\mathcal{F}_n\right] \leqslant C < \infty$ *a.s.* Indeed, by Tonelli's theorem one can show that for any $A \in \mathcal{F}_n$

$$\int_A \mathbf{E}\left[\sum_{k\geqslant 0} \frac{1}{(k+1)^3} |\Delta X_k| \, \mathbb{I}_{A(n-1,k)} \Big| \mathcal{F}_n\right] d\mathbb{P} \leqslant \int_A \sum_{k\geqslant 0} \frac{1}{(k+1)^2} d\mathbb{P},$$

As this inequality is for any set A, $\mathcal{F}_n-$measurable the claim holds. Then $\mathbf{E}\left[u_n \,|\mathcal{F}_n\right] = 0$ (by dominated convergence for conditional expectations). Hence, $M_n = \sum_{k=0}^n u_k$ is a martingale. We claim that we can write the generalised martingale as a martingale transform. Indeed, define the predictable process $\gamma_n = \sum_{k=0}^\infty (k+1)^3 \, \mathbb{I}_{A(n-1,k)}$

Since $\mathbf{E}\left[\Delta X_{n+1} \,|\mathcal{F}_n\right] = 0$ and $A(n-1,k)$ is $\mathcal{F}_n$ measurable, it follows that $\mathbf{E}\left[u_n \,|\mathcal{F}_n\right] = 0$, therefore $M_n = \sum_{k=1}^n u_n$ is a martingale.
Let $\gamma_n := \sum_{k=0}^\infty (k+1)^3 \, \mathbb{I}_{A(n-1,k)}$, we claim that X_n is a martingale transform and in fact $X_n = (\gamma \bullet M)_n$ [4]

$$\sum_{k=1}^n \langle \gamma_k, \Delta M_k \rangle = \sum_{k=1}^n \left(\sum_{i=0}^\infty (i+1)^3 \, \mathbb{I}_{A(k-1,i)}\right) \left(\sum_{j=0}^\infty \frac{1}{(j+1)^3} \Delta X_k \mathbb{I}_{A(k-1,j)}\right), \tag{2.3.4}$$

As the sets $A(k-1,j)$ are disjoint for fixed $m \geqslant 1$, the sums in (2.3.4) are reduced to

$$\sum_{k=1}^n \sum_{j=0}^\infty \Delta X_k \mathbb{I}_{A(k-1,j)} = \sum_{k=1}^n \Delta X_k = X_n - X_0. \tag{2.3.5}$$

Hence $X_k = X_0 + \sum_{k=1}^n \langle \gamma_k, \Delta M_k \rangle$.

□

In general, there are well-known conditions that allow to deduce when a local martingale is a martingale or a supermartingale. Indeed,

I A local martingale $\{X_i\}_{i\leqslant n}$ such that $\mathbf{E}\left(\sup_{i\leqslant n} (X_i)^-\right) < \infty$ for all $n \in \mathbb{N}$ is a supermartingale.

Indeed, let $Z = \sup_{i\leqslant n} (X_i)^-$ then $X_i + Z \geqslant 0$ for all $i \leqslant n$. Let $\{\tau_k\}_{k\geqslant 1}$ be a localizing sequence

[4] We denote the (local) martingale transform and the stochastic integral $(\gamma \bullet M)_n$

for X, then Fatou's lemma implies

$$\mathbf{E}\left[\liminf_k X_{\tau_k \wedge i} + Z\,|\mathcal{F}_{i-1}\right] \leqslant \liminf_k \mathbf{E}\left[X_{\tau_k \wedge i} + Z\,|\mathcal{F}_{i-1}\right] = \liminf_k X_{\tau_k \wedge (i-1)} + \mathbf{E}\left[Z\,|\mathcal{F}_{i-1}\right], \tag{2.3.6}$$
$$\mathbf{E}\left[X_i + Z\,|\mathcal{F}_{i-1}\right] \leqslant \liminf_k \mathbf{E}\left[X_{\tau_k \wedge i} + Z\,|\mathcal{F}_{i-1}\right] = X_{\tau_k \wedge (i-1)} + \mathbf{E}\left[Z\,|\mathcal{F}_{i-1}\right],$$

As $\mathbf{E}Z < \infty$ the r.v. $\mathbf{E}\left[Z\,|\mathcal{F}_{i-1}\right]$ is finite $\mathbb{P} - a.s.$ this implies that X is a supermartingale.

II A local martingale X such that $\mathbf{E}\sup_{i\leqslant n}|X_i| < \infty$ for all n is a martingale. The last fact follows as one can apply Lebesgue's dominated convergence theorem and obtain an equality in (2.3.6).

In fact, in the discrete-time setting, weaker assumptions yield the same result. This is Theorem 2.2 in [27]

Theorem 2.3.8. a) *Let X be a local martingale such that $\mathbf{E}X_n^- < \infty$ or $\mathbf{E}X_n^+ < \infty$ for all n then X is a martingale.*

b) *Let $\{X_i\}_{1\leqslant i\leqslant n}$ be a local martingale such that $\mathbf{E}X_N^- < \infty$ or $\mathbf{E}X_N^+ < \infty$ then X is a martingale.*

Proof. Let $\{X_n\}_{n\geqslant 0}$ be a local martingale and let $\{\tau_k\}_{k\geqslant 1}$ be a localizing sequence of stopping times of $\{X_n\}$, assume $\mathbf{E}X_n^- < \infty$ by Theorem 2.3.7 $X \in GM$, then $\mathbf{E}\left[X_{i+1}\,|\mathcal{F}_i\right] = X_i$ $\mathbb{P} - a.s.$ and $\mathbf{E}\left[|X_{i+1}|\,|\mathcal{F}_i\right] < \infty$ $\mathbb{P} - a.s.$ and the conditional Jensen's inequality implies $X_i^- \leqslant \mathbf{E}\left[X_{i+1}^-\,|\mathcal{F}_i\right]$ thus, $\mathbf{E}X_i^- \leqslant \mathbf{E}X_{i+1}^-$.
We claim that this implies $\mathbf{E}\left(X_i^+\right) < \infty$ for any $i \leqslant n$, indeed

$$\begin{aligned}\mathbf{E}X_i^+ = \mathbf{E}\left(\underline{\lim}_k X^+_{\tau_k \wedge i}\right) &\leqslant \underline{\lim}_k \mathbf{E}X^+_{\tau_k \wedge i} = \mathbf{E}X_{\tau_k \wedge i} + \mathbf{E}X^-_{\tau_k \wedge i}\\ &\leqslant \underline{\lim}_n \left\{\mathbf{E}X_0 + \sum_{i=0}^n \mathbf{E}X_i^-\right\} < \infty,\end{aligned}$$

then $\mathbf{E}|X_i| < \infty$ for any $i \leqslant n$ and taking into account that $X \in GM$ we have that X is a martingale, as n is any positive integer, both claims $a)$ and $b)$ have been proved. □

Finally, we have a result relating an integrability condition on the terminal value of a local martingale and the martingale property, this yields the same conclusion as Theorem 2.3.8. This was pointed out in [20], which we follow closely

Theorem 2.3.9. *Suppose $\{X_n\}_{1\leqslant n\leqslant N}$ is a real-valued local martingale and there is a constant $M \in \mathbb{R}$ such that $X_N \geqslant M$ then $\{X_n\}_{1\leqslant n\leqslant N}$ is a martingale.*

Proof. If $M > 0$ then, this is a trivial consequence of Theorem 2.3.8. Otherwise, let $\{\tau_k\}_{k\geqslant 1}$ be a localising sequence for X. Theorem 2.3.6 implies that X is a generalised martingale then

$$\mathbf{E}\left[X_N|\mathcal{F}_n\right] = X_n\ \mathbb{P} - a.s. \tag{2.3.7}$$

thus $X_n \geqslant M$ and then $\mathbf{E}X_n^- < \infty$. This in turn implies that for any $k > 0$, $X_{\tau_k \wedge n} \geqslant M$. By Fatou's lemma,

$$\mathbf{E}X_0 = \lim_{k\to\infty} \mathbf{E}X_{\tau_k \wedge 0} = \lim_{k\to\infty} \mathbf{E}X_{\tau_k \wedge n} \geqslant \mathbf{E}\big(\lim_{k\to\infty} X_{\tau_k \wedge n}\big) = \mathbf{E}X_n. \tag{2.3.8}$$

As $\mathbf{E}X_n \geqslant \mathbf{E}X_n^+$ this implies that $\mathbf{E}|X_n| < \infty$. Thus $\{X_n\}_{n\geqslant 1}$ is a martingale. □

Notice that in a continuous-time setting, we only obtain an inequality in (2.3.7) and the condition $X_t \geqslant M$ only yields that X_t is a supermartingale.

4 Discrete-time optimal investment

4.1 Setting and preliminary considerations

In this section we shall describe a general incomplete market model in discrete time, we shall explain briefly the main ideas on important results that justify the existence of a martingale measure under with desirable integrability properties. These properties are used in Chapter 3. We consider the usual probabilistic setting, namely, a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ and $\mathcal{F}_0$ is complete with respect to $\mathbb{P}$ and a $d-$dimensional adapted stochastic process $(S_t)_{t=0,1,\ldots,T}$ describing the evolution of the prices of d risky assets in a given economy.
For simplicity, we assume that the market is frictionless, there are no costs involved in trading or borrowing.
The investor creates a portfolio by trading assets in the market with an initial capital or initial endowment x. The market consists of d risky assets that we call stocks or shares and there is a riskless asset, denoted by S_t^0, in addition, we always assume that the amount of shares invested at time t are decided at time $t-1$. In other words, we may interpret this riskless asset as the money market (a bond or a bank account) and the processes $\left(S_t^i\right)_{t=0,1,\ldots,T}$ with $i = 1, \ldots, d$ will be the stocks prices at times $t = 0, 1, \ldots, T$. Here $\langle \cdot, \cdot \rangle$ denotes the 'Euclidean' inner product in $\mathbb{R}^d$, i.e. $\langle x, y\rangle = \sum_{i=1}^d x_i y_i$ and $\Delta S_i = S_i - S_{i-1}$. We denote by φ^0 and $\varphi \in \mathbb{R}^d$, the holdings in the riskless asset and in the risky asset, respectively.
Moreover, we shall assume that investors always trade in a self-financing way.[5] The same condition at time 0 is equivalent to have the following 'budget constraint' $x = \varphi_1^0 \cdot S_0^0 + \left\langle \varphi^1, S_0\right\rangle$, we assume that $S_0^0 = 1$. If the $\mathbb{R}^{d+1}$-valued process S_t represents the asset prices at time t and we take the riskless asset as a discount factor, the discounted asset prices are denoted by $\tilde{S}_t^i := S_t^i / S_t^0$.
If holdings on assets are self-financing, then the change in the value of the portfolio are only due to changes in the values of the asset prices. A self-financing strategy $\{\varphi_t\}_{t\geqslant 1}$ satisfies a similar condition when asset prices are discounted, namely $\varphi_{t+1}^0 + \left\langle \varphi_{t+1}, \tilde{S}_t\right\rangle = \varphi_t^0 + \left\langle \varphi_t, \tilde{S}_t\right\rangle$ and in the case $t = 0$ the self-financing condition is $\varphi_1^0 + \left\langle \varphi_1, \tilde{S}_0\right\rangle = x$ therefore

$$\Delta\left(\varphi_t^0 + \left\langle \varphi_t, \tilde{S}_t\right\rangle\right) = \sum_{i=0}^d \varphi_t^i \tilde{S}_t^i - \sum_{i=0}^d \varphi_{t-1}^i \tilde{S}_{t-1}^i = \varphi_t^0 - \varphi_{t-1}^0 + \left\langle \varphi_t, \tilde{S}_t\right\rangle - \left\langle \varphi_{t-1}, \tilde{S}_{t-1}\right\rangle,$$

$$\Delta\left(\varphi_t^0 + \left\langle \varphi_t, \tilde{S}_t\right\rangle\right) = \left\langle \varphi_t, \tilde{S}_t\right\rangle - \left\langle \varphi_t, \tilde{S}_{t-1}\right\rangle = \left\langle \varphi_t, \Delta\tilde{S}_t\right\rangle = \sum_{i=1}^d \varphi_t^i \Delta\tilde{S}_t^i, \tag{2.4.1}$$

[5] We say a trading strategy $\left\{\varphi_t^0, \varphi_t^1, \ldots, \varphi_t^d\right\}_{t\leqslant T}$ is a *self-financing* trading strategy on $\left(S_t^i\right)_{i=1,\ldots,d}$ if $\varphi_t^0 S_t^0 + \langle \varphi_t, S_t\rangle = \varphi_{t+1}^0 S_t^0 + \langle \varphi_{t+1}, S_t\rangle$ for any $t \leqslant T$. In other words, no extra capital is injected or extracted from the portfolio, at any time. A strategy is said to be *predictable* if the holdings at time t, are $\mathcal{F}_{t-1}-$measurable i.e. for all t we have $\varphi_t \in \mathcal{F}_{t-1}$.

by the self-financing condition. Adding up, the equalities in (2.4.1), adding $\langle \varphi_1, \Delta\tilde{S}_1 \rangle$ and using the self-financing condition (or the budget constraint)

$$\langle \varphi_t, \tilde{S}_t \rangle = x + \sum_{k=1}^{t} \langle \varphi_k, \Delta\tilde{S}_k \rangle .$$

The process $\tilde{V}_t^{x,\varphi}$ represents the value of the holdings at time t, starting with an initial investment x, it is referred as the *portfolio value process*. A trading strategy on the market can be represented by a $d-$dimensional vector $\{\varphi_t\}_{t\geqslant 1}$ describing the holdings on the risky assets at time t, we emphasise the assumption on the trading strategies, namely, they are $\mathcal{F}_t$-predictable, and the initial investment x. Denote by $\mathcal{P}$ the set of all possible holdings on the d assets throughout the periods $1, \ldots, T$.
Considering discounted asset prices, a trading strategy consists of two components, the process $\{\phi_k\}_{k=1,\ldots,T}$ and the initial investment x, this represents the initial value of the portfolio at time 0, and the investor's discounted wealth evolves according to the process given by

$$\tilde{V}_t^{x,\phi} = x + \sum_{k=1}^{t} \langle \phi_k, \Delta\tilde{S}_k \rangle . \qquad (2.4.2)$$

Proposition 2.4.1. *Suppose* (V_0, ϕ_t) *is such that* $\phi \in \mathcal{P}$, $V_0 \in \mathbb{R}$ *is* $\mathcal{F}_0-$*measurable then there is a predictable self-financing strategy* (φ_t^0, ϕ_t) *such that* $V_t = \varphi_t^0 S_t^0 + \langle \phi_t, S_t \rangle$ *or equivalently,* $\tilde{V}_t = \varphi_t^0 + \langle \phi_t, \Delta\tilde{S}_t \rangle$.

Proof. Indeed, by the definition of the wealth process we should have $V_t = \varphi_t^0 S_t^0 + \langle \phi_t, S_t \rangle$, on the other hand, by the self-financing condition, $\tilde{V}_t = \tilde{V}_{t-1} + \langle \phi_t, \Delta\tilde{S}_t \rangle$ for all $t = 1, \ldots, T$, as φ_t^0 should be such that or $\tilde{V}_t = \varphi_t^0 + \langle \phi_t, \tilde{S}_t \rangle$ define inductively, $\varphi_1^0 := V_0 + \langle \phi_1, -\tilde{S}_0 \rangle$, and

$$\varphi_t^0 := \tilde{V}_{t-1} + \langle \phi_t, -\tilde{S}_{t-1} \rangle .$$

$\square$

By Proposition 2.4.1, hereafter, we shall work with discounted prices and trading strategies on the risky assets. In order to simplify our notation, we shall write S_t instead of $\tilde{S}_t$, similarly the discounted wealth process $\tilde{V}_t^{x,\phi}$ will be denoted by $V_t^{x,\phi}$.
Therefore, (2.4.2) is written as

$$V_t^{x,\phi} = x + \sum_{k=1}^{t} \langle \phi_k, \Delta S_k \rangle . \qquad (2.4.3)$$

We now describe important definitions related to our models. We denote by $D_t(\omega)$ the affine subspace of the support of the regular conditional distribution of ΔS_t given $\mathcal{F}_{t-1}$.
We denote by Ξ_t, the set of $\mathcal{F}_t$-measurable $d-$dimensional random variables. We first define an arbitrage opportunity in the market. This concept describes the idea of an investment with no cost at the initial time that generates possibly a positive gain with no risk (non-negative wealth)

Definition 2.4.2. An arbitrage opportunity is a trading strategy $\phi \in \mathcal{P}$ such that

- $V_T^{0,\phi} \geqslant 0, \ \mathbb{P} - a.s.$

- $\mathbb{P}\left(V_T^{0,\phi} > 0\right) > 0$.

Definition 2.4.3. We say that the market (described by the filtered probability space together with the process S) has no arbitrage opportunity if for any strategy ϕ such that

$$V_T^{0,\phi} \geqslant 0 \ \mathbb{P} - \text{a.s. implies } \ V_T^{0,\phi} = 0 \ \mathbb{P} - \text{a.s.} \tag{2.4.4}$$

Definition 2.4.4. We say that the market is **complete** if every bounded $\mathcal{F}_T$−measurable random variable F (discounted price) can be hedged by a self-financing strategy, i.e. there is a self-financing ϕ and an initial investment x such that

$$V_T^{x,\phi} = F. \tag{2.4.5}$$

Whenever a market is not complete (not all the contingent claims can be hedged by means of a replicating portfolio), we say that the market is **incomplete**.

Definition 2.4.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathbb{Q}$ be a probability measure on $\mathcal{F}$. We say $\mathbb{Q}$ is an *equivalent martingale measure* if

- $\mathbb{Q}$ is equivalent to $\mathbb{P}$ [6]
- The process $(S_n, \mathcal{F}_n)_{n=0,\dots,T}$ is a martingale with respect to $\mathbb{Q}$.

The main theorem relating Definitions 2.4.3 and 2.4.5 is the following

Theorem 2.4.6. *Refer to [16] The following statements are equivalent:*

- *There exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that under $\mathbb{Q}$ the process $(S_n, \mathcal{F}_n)_{n=0,\dots,T}$ is a martingale.*
- *There are no arbitrage opportunities in the market.*

A general result was proved in a more general setting by Delbaen and Schachermayer in [18] and [19].

Theorem 2.4.7. *Assume that the market is arbitrage free, the market is complete if and only if the set of equivalent martingale measures is a singleton.*

The next proposition yields a useful quantitative characterisation of a market having no arbitrage, namely, at each time $t = 1, \dots, T$ and for trading strategies $\{\phi_t\}_{t\geqslant 1}$ such that $\phi_t \in D_t(\omega)$ (see paragraph before Definition 2.4.2) for all $t \geqslant 1$ the probability that the value of the portfolio may be less than a strictly negative value $-\beta_t$ is strictly positive.
To explain such a characterization, we define the following set of trading strategies, we need to avoid the cases when ϕ_t is orthogonal $\mathbb{P}$−a.s to ΔS_t.

Definition 2.4.8. Define the set of strategies $\overline{\Xi}_t$ as

$$\overline{\Xi}_t := \{\theta \in \Xi_t \, : \, |\theta(\omega)| = 1 \text{ on } \{\omega : \, D_t(\omega) \neq \{0\}\}, \theta \in D_t(\omega), \mathbb{P} - \text{a.s.}\} \tag{2.4.6}$$

[6] Two measures are said to be equivalent if they have the same sets of measure zero.

Theorem 2.4.9. *Suppose the market has no arbitrage, i.e. (2.4.4) holds, then there exist* $\mathcal{F}_t-$*measurable random variables* β_t, $\kappa_t > 0$ *satisfying*

$$\operatorname*{ess\,inf}_{p\in\overline{\Xi}_t} \mathbb{P}\left(\langle p, \Delta S_{t+1}\rangle < -\beta_t \,|\mathcal{F}_t\right) > \kappa_t \tag{2.4.7}$$

for all $t = 0, 1, \ldots, T$, $\mathbb{P} - a.s.$

Proof. Fix $t \in \{1, \ldots, T\}$ and let $\{\delta_n\}_{n\geqslant 0}$ be a decreasing sequence tending to 0. Define

$$A_n := \left\{\omega : \operatorname*{ess\,inf}_{p\in\overline{\Xi}_t} \mathbb{P}\left(\langle p, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right) = 0\right\}.$$

The essential infimum is attained in the condition defining A_n. Indeed, by the properties of the essential infimum of a collection of random variables (see [38], p) it follows that there is a sequence of random variables $\{p_k^n\}_{k\geqslant 1} \subset \overline{\Xi}_t$ such that

$$\mathbb{P}\left(\langle p_k^n, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right) - \frac{1}{k} < \operatorname*{ess\,inf}_{p\in\overline{\Xi}_t} \mathbb{P}\left(\langle p, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right).$$

By a compactness principle in $L^0(\Omega, \mathcal{F}, \mathbb{P})$[7] there is a subsequence $\tilde{p}_k^n$ converging to an $\mathcal{F}_t-$measurable random variable p^n. Let $B_n^k := \{\omega : \langle \tilde{p}_k^n, \Delta S_{t+1}\rangle < -\delta_n\}$, then $\liminf_k \mathbb{I}_{B_n^k} = \mathbb{I}_{\liminf_k B_n^k}$, this follows from $\inf_{k\geqslant m} \mathbb{I}_{B_n^k} = \mathbb{I}_{\cap_{k\geqslant m} B_n^k}$ and the continuity property of probability measures. Fatou's lemma implies

$$\mathbf{E}\left[\liminf_k \mathbb{I}_{B_n^k} \,|\mathcal{F}_t\right] \leqslant \liminf_k \mathbb{P}\left(\langle \tilde{p}_k^n, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right). \tag{2.4.8}$$

Let $B_n = \{\omega : \langle p^n, \Delta S_{t+1}\rangle < -\delta_n\}$. Clearly, $B_n \subset \liminf_k B_n^k$ then from (2.4.8)

$$\mathbb{P}\left(\langle p^n, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right) \leqslant \operatorname*{ess\,inf}_{p\in\overline{\Xi}} \mathbb{P}\left(\langle p, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right),$$

The strategy p^n attains the essential infimum and we can actually express the set A_n as the set $\{\omega : \mathbb{P}\left(\langle p^n, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right) = 0\}$. Define $A := \bigcap_{n=1}^{\infty} A_n$, we claim that $\mathbb{P}(A) = 0$. Suppose that $\mathbb{P}(A) > 0$ then, by the same principle, we can take $\{p^n\}$ such that $p^n \to \tilde{p}$ $\mathbb{P} - a.s.$ and a similar argument, we have $\mathbb{P}\left(\langle \tilde{p}, \Delta S_{t+1}\rangle < 0 \;\; |\mathcal{F}_t\right) \leqslant \underline{\lim}_n \mathbb{P}\left(\langle p^n, \Delta S_{t+1}\rangle < -\delta_n \,|\mathcal{F}_t\right) = 0$ *on* A. Then $\mathbb{P}\left(\langle \tilde{p}, \Delta S_{t+1}\rangle \geqslant 0 \,|\mathcal{F}_t\right) = 1$ therefore $\mathbb{P}\left(\langle \tilde{p}, \Delta S_{t+1}\rangle \geqslant 0\right) = 1$. Let $\tilde{\phi}_i = 0$ if $i \neq t+1$ and $\tilde{\phi}_t = \tilde{p} \cdot \mathbb{I}_A$ and $V^{0,\tilde{\phi}} = \left\langle \tilde{\phi}, \Delta S_{t+1}\right\rangle$ as the market has no arbitrage it follows that $\mathbb{P}\left(\left\langle \tilde{\phi}, \Delta S_{t+1}\right\rangle = 0 \,|\mathcal{F}_t\right) = 1$ $\mathbb{P} - a.s.$. This contradicts our assumption, as we have $\tilde{p} \in D_{t+1}(\omega)$. The random variable $\tilde{p}$ is a limit of elements in $D_{t+1}(\omega)$, as this set is actually a closed subset then $\tilde{p} \in D_{t+1}(\omega)$. On the other hand, $\tilde{V} = \left\{x \in \mathbb{R}^d : \langle \tilde{p}\mathbb{I}_A, \Delta S_{t+1} > 0\rangle\right\}$ is an open neighbourhood of $\tilde{p}\mathbb{I}_A$ whose measure under $\mu(\cdot, \omega) = \mathbb{P}\left(\Delta S_{t+1} \in \cdot \,|\mathcal{F}_t\right)$ is zero, this contradicts $\tilde{p} \in D_{t+1}(\omega)$, (see Definition of 2.4.8) then $\mathbb{P}(A) = 0$. Define

$$\beta_t := \sum_{i=1}^{\infty} \delta_i \mathbb{I}_{A_i^C / A_{i-1}^C},$$

Notice that all the sets A_i depend on t and such sets are $\mathcal{F}_t-$measurable.

It follows from the previous paragraph that $\mathbb{P}(\Omega/A) = 1$ and $A_n^c \subset A_{n+1}^c$ for all $n \geqslant 1$ and

[7] See appendix B

$\sum_{n=1}^{\infty}\mathbb{I}_{A_n^c/A_{n-1}^c}=\mathbb{I}_{A^c}=1\ \mathbb{P}-a.s.$.
We define the random variable $\kappa_t>0$ on Ω/A. Indeed, if $\omega\in A_n^c/A_{n-1}^c$ then

$$\operatorname*{ess\,inf}_{p\in\Xi_t}\mathbb{P}\left(\langle p,\Delta S_{t+1}\rangle<-\beta_t\,|\mathcal{F}_t\right)=\mathbf{E}\left[\mathbb{I}_{\left\{\langle\tilde{p},\Delta S_{t+1}\rangle<-\beta_t\right\}}\mathbb{I}_{A_n^c/A_{n-1}^c}\,|\mathcal{F}_t\right]$$

$$=\mathbf{E}\left[\mathbb{I}_{\left\{\langle\tilde{p},\Delta S_{t+1}\rangle<-\delta_n\right\}}\mathbb{I}_{A_n^c/A_{n-1}^c}\,|\mathcal{F}_t\right]=\mathbb{P}\left(\langle\tilde{p},\Delta S_{t+1}\rangle<-\delta_n\,|\mathcal{F}_t\right)>0,$$

as previously discussed, the essential infimum is attained by a random variable $\tilde{p}$, and by definition of A_n and the fact that $\beta_t=\delta_n$ on A_n^c.
Recall that $\kappa_t:=\operatorname{ess\,inf}_{p\in\tilde{\Xi}_t}\mathbb{P}\left(\left\langle\tilde{\phi},\Delta S_{t+1}\right\rangle<-\beta_t\,|\mathcal{F}_t\right)$ is $\mathcal{F}_t-$measurable[5]. The lemma is proved. □

4.2 Equivalent martingale measures and utility maximisation

In the following section we shall describe results in utility maximisation in discrete-time that are crucial to prove some estimates connected to the absence of arbitrage. Utility functions are used to describe economic agents' preferences, the problem of utility maximisation has a straightforward interpretation in economy and finance, and it has been studied extensively.
The main justification of its use comes from economic theory, preferences and attitudes towards risky outcomes can be described by means of utility functions. Moreover, we shall see an important connection of this problem with Theorem 2.4.6 and the existence of martingale measures for S.
The following assumptions are made throughout this section.

Assumption 2.4.1. *The utility function $U:\mathbb{R}\to\mathbb{R}$ is a concave and non-decreasing function and continuously differentiable; $U(0)=0$ and there is $\tilde{x}>0$ and $\gamma\in(0,1)$ such that for $x\geqslant\tilde{x}$ and any $\lambda\geqslant 1$*

$$U(\lambda x)\leqslant\lambda^{\gamma}U(x)\,. \tag{2.4.9}$$

And there exist $x_{-}<0$ and $\alpha>0$ such that for any $\lambda\geqslant 1$ and $x\leqslant x_{-}$

$$U(\lambda x)\leqslant\lambda^{1+\alpha}U(x)\,. \tag{2.4.10}$$

We consider the problem of maximising a utility function of the terminal wealth with initial capital x, with U a function satisfying Assumption 2.4.1 . In other words, this problem is concerned with the computation of the trading strategy (if such a strategy exists) such that

$$u(x):=\sup_{\phi\in\Phi(x)}\mathbf{E}U\left(V_T^{x,\phi}\right)=\mathbf{E}U\left(V_T^{x,\phi^{\star}}\right),$$

where $\Phi(x)$ is the class of $\mathcal{F}_t-$predictable processes with

$$\mathbf{E}U\left(V_T^{x,\phi}\right)=\mathbf{E}U\left(x+\sum_{i=1}^{d}\langle\phi_i,\Delta S_i\rangle\right)<\infty. \tag{2.4.11}$$

The function $u(x)$ denotes the maximum 'satisfaction' an investor can obtain by financing his portfolio with an initial amount x.
The approach in [44] to ensure the existence of an optimal strategy is a traditional one, namely, the method of dynamic programming. It must be pointed out that while backward induction

yields concrete solutions to optimisation problems when the underlying process is Markov or a deterministic function, we do not assume any Markov property whatsoever.
As investors assess their trading strategies by means of the utility function $U(x)$ this naturally leads to consider (random) functions depending jointly on x and ω. For instance, suppose that investor starts trading at time $t < T$, with an initial capital x and $\{\xi_s(\omega)\}_{s>t}$ in the risky assets then the expected 'satisfaction' on this trading at time $t+1$ is equal to $\mathcal{U}(x,\omega) = \mathbf{E}\left[U\left(x + \langle \xi_t, \Delta S_{t+1}\rangle\right) \middle| \mathcal{F}_t\right]$, as the function U is concave by linearity of the conditional expectation, it follows that, for fixed $\omega \in \Omega$, the function $\mathcal{U}_t(x,\omega)$ is concave in $\mathbb{R}$, hence is continuous (for fixed ω). It is also $\mathcal{F}_t-$measurable thus it is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_t$-measurable.
In order to deal with computations involving random functions we make the following remark.

Remark 2.4.10. Assume (as it is in [44]) that we have a function $V : \Omega \times \mathbb{R}^d \to \mathbb{R}$. Such that for fixed $\omega \in \Omega$ the function V is continuous, by separability of $\mathbb{R}$, for each $\omega \in \Omega$, the function $V(\omega,\cdot)$ is determined at the values on a dense countable set. Thus, without loss of generality, when we find it convenient, we may identify the function V with a random variable taking values in $\varGamma := \mathbb{R}^{\mathbb{N}}$.

Assumption 2.4.1 implies an estimate of 'power' or 'polynomial' type (with respect to λ). This is expected as this type of result holds for some intervals of the form $[\tilde{x}, \infty)$ and $(-\infty, x_-]$ and U is a concave function on $\mathbb{R}$.

Proposition 2.4.11. *There are* $0 < \gamma < 1$ *and* $C > 0$*, constants such that for any* $x \in \mathbb{R}$ *and* $\lambda \geqslant 1$

$$U(\lambda x) \leqslant \lambda^\gamma U(x) + \lambda^\gamma C, \tag{2.4.12}$$

$$U(\lambda x) \leqslant \lambda U(x) + \lambda^\gamma C. \tag{2.4.13}$$

Proof. Suppose that $x > \tilde{x}$, then (2.4.12) follows from Assumption 2.4.1 for any $C > 0$, as mentioned above, $U(0) \leqslant U(x)$ then (2.4.13) follows from (2.4.12).
Suppose that $x < 0$. We prove (2.4.13) first. By concavity, the function U' is decreasing, the mean value theorem implies that $xU'(x) \leqslant U(x) - U(0)$ then

$$U(\lambda x) \leqslant U(x) + U'(x)\,x(\lambda - 1) \leqslant U(x) + (U(x) - U(0))(\lambda - 1), \tag{2.4.14}$$

$$U(\lambda x) \leqslant \lambda U(x) + \lambda^\gamma C, \tag{2.4.15}$$

for any $C > 0$. The second inequality follows similarly, from the fact that $U(x) \leqslant 0$ and $\lambda \geqslant 1$ and $0 < \gamma < 1$ in (2.4.14) yields $(U(x) - U(0))(\lambda^\gamma - 1) \geqslant (U(x) - U(0))(\lambda - 1)$ then

$$U(\lambda x) \leqslant U(x) + U'(x)\,x(\lambda^\gamma - 1) \leqslant U(x) + (U(x) - U(0))(\lambda^\gamma - 1) = \lambda^\gamma U(x) + \lambda^\gamma C, \tag{2.4.16}$$

for any $C > 0$. If $0 \leqslant x \leqslant \tilde{x}$ then $U(\lambda x) \leqslant \lambda^\gamma U(\lambda\tilde{x})$, as $\lambda^\gamma \geqslant 1$. Finally $U(\lambda x) \leqslant \lambda^\gamma C + \lambda^\gamma U(x)$ and the last expression is majorised by $\lambda^\gamma C + \lambda U(x)$. □

Clearly if $U(x) > 0$ for all $x \in \mathbb{R}$ then (2.4.12) implies (2.4.13). Thus, (2.4.13) helps to manage the case when $U(x)$ may take negative values. Moreover, while the conditions in Assumption 2.4.1 describe the asymptotic behaviour of the function U at $\pm\infty$, (2.4.12) and (2.4.13) in Proposition 2.4.11 hold for any $x \in \mathbb{R}$.

One can have these inequalities to be valid for functions that depend on other variables, for instance, when the function U in Proposition 2.4.11 depends on $\omega \in \Omega$ as well.
Combining this consideration and the ideas underlying the dynamic programing principle, leads to the definitions of the 'value' functions of the problem of utility maximisation. See Assumption 2.4.2 below.

We define the functions $\{U_t(x) : t = 0, 1, \ldots, T\}$ and we shall make the following assumption about them.

Assumption 2.4.2. *Suppose that the following random functions are well-defined*

$$U_T(x) := U(x), \text{ and } U_t(x) := \operatorname*{ess\,sup}_{\xi \in \Xi_t} \mathbf{E}\left[U_{t+1}(x + \langle \xi, \Delta S_{t+1} \rangle) | \mathcal{F}_t\right] \ x \in \mathbb{R}, \tag{2.4.17}$$

and for $t \in \{0, 1, \ldots, T\}$ and all $x \in \mathbb{R}$,

$$U_t(x) := \operatorname*{ess\,sup}_{\xi \in \Xi_t} \mathbb{E}\left[U_{t+1}(x + \langle \xi, \Delta S_{t+1} \rangle) | \mathcal{F}_t\right] < \infty \ a.s. \tag{2.4.18}$$

Further, suppose that for all $x \in \mathbb{R}$

$$\mathbf{E}U_0(x) < \infty. \tag{2.4.19}$$

Notice that for $t \in \{0, 1, \ldots, T-1\}$ such functions in (2.4.17) and (2.4.18) depend on ω. One could regard such functions as the optimal future wealth from t to $t+1$ if investment starts at time t with an initial capital x.

Proposition 2.4.12. *Suppose that Assumption 2.4.1 holds. Then the functions $U_t(x)$ satisfy (2.4.12) and (2.4.13) $\mathbb{P}-a.s.$*

Proof. This holds for $t = T$ by assumption and Proposition 2.4.11. Suppose that (2.4.12) and (2.4.13) hold for U_{t+1}. Using Lemma 4.9 and Proposition 4.10 in [44] we have

$$U_t(x) = \operatorname*{ess\,sup}_{\xi \in \Xi_t} \mathbf{E}\left[U_{t+1}(x + \langle \xi, \Delta S_{t+1} \rangle) | \mathcal{F}_t\right] = \mathbf{E}\left[U_{t+1}\left(x + \left\langle \tilde{\xi}(x), \Delta S_{t+1} \right\rangle\right) | \mathcal{F}_t\right], \tag{2.4.20}$$

for some $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_t-$measurable function $\tilde{\xi}(x, \omega)$. Then

$$U_t(\lambda x) = \mathbf{E}\left[U_{t+1}\left(\lambda x + \left\langle \tilde{\xi}(\lambda x), \Delta S_{t+1} \right\rangle\right) \Big| \mathcal{F}_t\right] \leqslant \mathbf{E}\left[U_{t+1}\left(\lambda\left(x + \left\langle \frac{\tilde{\xi}(\lambda x)}{\lambda}, \Delta S_{t+1} \right\rangle\right)\right) \Big| \mathcal{F}_t\right],$$

$$U_t(\lambda x) \leqslant \lambda^\gamma \operatorname*{ess\,sup}_{\xi \in \Xi_t} \mathbf{E}\left[\lambda^\gamma U_{t+1}\left(x + \left\langle \xi, \Delta S_{t+1} \right\rangle\right) + \lambda^\gamma C \Big| \mathcal{F}_t\right] \leqslant \lambda^\gamma U_t(x) + \lambda^\gamma C. \tag{2.4.21}$$

The inequality (2.4.13) is deduced similarly. ☐

Remark 2.4.13. If U is bounded from above then Assumption 2.4.2 holds.

Absence of arbitrage (Definition 2.4.3) is not only relevant as a desirable component in our model from an economic perspective, but in some cases is indispensable to ensure existence of optimal strategies.

Proposition 2.4.14. *If U is strictly increasing and there is an arbitrage opportunity in the market, then there is no trading strategy that attains the maximum.*

Proof. If $\hat{\phi}$ is an arbitrage opportunity and suppose that $\phi^{\star}$ is an optimiser, then

$$\mathbf{E}U\left(V_T^{x,\phi^{\star}+\hat{\phi}}\right) = \mathbf{E}U\left(V_T^{x,\phi^{\star}} + V_T^{0,\hat{\phi}}\right) > \mathbf{E}U\left(V_T^{x,\phi^{\star}}\right). \tag{2.4.22}$$

□

The following result is crucial for chapter 3. We follow the ideas developed in [44].

Definition 2.4.15. Define the class of random variables $\mathscr{W}$ to be the set of random variables with finite moments of all orders.

We shall explain in more detail section 7 in [44].

Theorem 2.4.16. *Suppose that U is a non-decreasing concave function that satisfies Assumption 2.4.1. Further, suppose that for some k, l, M and $K \geqslant 0$*

$$M\,|x|^{-l} \leqslant U'(x) \leqslant K\left(|x|^k + 1\right), \text{ for all } x \in \mathbb{R}. \tag{2.4.23}$$

Suppose in addition that U is continuously differentiable and strictly increasing.
Furthermore, assume that for all $t \in \{0, 1, \ldots, T\}$, we have $|\Delta S_t| \in \mathscr{W}$ and that the condition 2.4.4 holds and the processes β_t and κ_t in Lemma 2.4.9 are such that $\frac{1}{\beta_t}, \frac{1}{\kappa_t} \in \mathscr{W}$ for all $t \in \{0, 1 \ldots, T-1\}$. Then under these conditions it follows that

- *The 'value' functions U_t are well-defined.*
- *For all $x \in \mathbb{R}$ it holds $U_t(x) < \infty$ $\mathbb{P} - a.s.$ and for all $x \in \mathbb{R}$*

$$\mathbf{E}U_0(x) < \infty. \tag{2.4.24}$$

For every initial endowment x there is a trading strategy $\phi^{\star}(x) \in \Phi(x)$ such that

$$u(x) = \sup_{\phi \in \Phi(x)} \mathbf{E}U\left(V_T^{x,\phi}\right) = \mathbf{E}U\left(V_T^{x,\phi^{\star}}\right). \tag{2.4.25}$$

Theorem 2.4.16 and the existence of the strategy $\phi^{\star}$ allows to obtain a martingale measure, for S, given by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \frac{U'\left(V_T^{x,\phi^{\star}}\right)}{\mathbf{E}U'\left(V_T^{x,\phi^{\star}}\right)} \tag{2.4.26}$$

Remark 2.4.17. By choosing a suitable utility function, one can ensure the existence of a martingale measure $(\mathbb{Q} \sim \mathbb{P})$ and whose density $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is bounded below. (For instance, letting $U(x) := x$ if $x > 0$ and $U(x) = U_1(x)$ otherwise, such that U is continuously differentiable and concave in $\mathbb{R}$ and satisfies Assumption 2.4.1) one can see that due to (2.4.23)

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{U'\left(V_T^{x,\phi^{\star}}\right)}{\mathbf{E}U'\left(V_T^{x,\phi^{\star}}\right)} \geqslant \frac{1}{\mathbf{E}U'\left(V_T^{x,\phi^{\star}}\right)} > 0.$$

On the other hand $\mathbf{E}U'\left(V_T^{x,\phi^{\star}}\right) < \infty$, as we will explain, the proof of Theorem 2.4.16 shows that $\mathbf{E}\,|\phi_t^{\star}|^{\alpha} < \infty$, for any $\alpha > 0$ (in fact $|\phi_t^{\star}| \leqslant \psi_t\left(1 + |x|^{r_t}\right)$ for a positive constant r_t and a random variable $\psi_t \in \mathscr{W}$, together with the assumption on U', condition (2.4.23) implies that $0 < \mathbf{E}U'\left(V^{x,\phi^{\star}}\right) \leqslant K\left(1 + \mathbf{E}\left|V_T^{x,\phi^{\star}}\right|^k\right) < \infty$ as $\Delta S_{t+1} \in \mathscr{W}$. Under models satisfying the assumptions of Theorem 2.4.16 the class of martingale measures is 'rich enough' to find a density that is bounded below and $\frac{d\mathbb{Q}}{d\mathbb{P}} \in \mathscr{W}$.

4.3 Remarks on Theorem 2.4.16

The proof of Theorem 2.4.16 is beyond the scope of this work, but we shall make some remarks and explain in detail the main ideas of section 7 in [44].
To simplify our notation, we shall omit the dependence on ω as it is clear that $\{U_t(x)\}_{t=0,1,\ldots,T}$ are $\mathcal{F}_t$-measurable random variables.
Backward induction is applied in order to obtain the statements for all $t \leqslant T$, as explained above the functions $U_T(x)$ and $\{U_t(x)\}_{t=0,\ldots,T}$ defined in (2.4.17) and (2.4.18) specify the optimal value in one-step period. It turns out that, by our assumptions on U' and U in Theorem 2.4.16, one can show that such 'value functions' $U_t(x)$ are dominated by random variables that are defined in terms of U and a polynomial function of the initial cost x. In other words, for every $t \leqslant T-1$ there is a constant $r_t > 0$ and a random variable $\rho_t \in \mathscr{W}$ such that $\mathbb{P} - a.s.$

$$U_t(x,\omega) \leqslant \mathbf{E}\left[U\left(|x| + (1+|x|)^{r_t}\rho_t\right)|\mathcal{F}_t\right]. \tag{2.4.27}$$

Notice that (2.4.27) implies that for all $t \in \{0,1,\ldots,T-1\}$,

$$U_t(x) < \infty \quad a.s. \tag{2.4.28}$$

Furthermore, Lemma 4.5 in [44] states that the 'essential infimum is attained', thus applying this lemma to $V(x,\omega) := U_{t+1}(\omega, x + \langle\theta, \Delta S_{t+1}(\omega)\rangle)$ implies that there is a random variable $\tilde{\xi}_t(x)$ (depending on x and being measurable) such that the following equality holds

$$U_t(x) = \mathbf{E}\left[U_{t+1}\left(x + \left\langle\tilde{\xi}_t(x), \Delta S_t\right\rangle\right)|\mathcal{F}_t\right],$$

and it ensures that if $U_{t+1}(x)$ is continuously differentiable and concave then so it is $U_t(x)$. As $U_T(x) = U(x)$ all the 'one-step' value functions in (2.4.18) are also continuously differentiable (for fixed ω).
Moreover, fixing the initial portfolio's capital x, the proof of Proposition 7.1 in [44] shows that trading strategies taking 'large values' does not yield higher utility, thus, a strategy $\tilde{\xi}_t(x)$ should be optimal only if for every $t = 1, 2, \ldots T$ there are random variables $\psi_t \in \mathscr{W}$ and constants $\zeta_t > 0$ such that

$$\left|\tilde{\xi}_t(x)\right| \leqslant \psi_t\left(1 + |x|^{\zeta_t}\right). \tag{2.4.29}$$

In this section we explain why this holds. In the case $t = T-1$ (2.4.27) is obvious, we can take $\rho_T = 0$, as $U_T = U$. By Proposition 5.2 in [44] (this is actually a consequence of a 'relatively compactness' principle[8] in $L^0(\Omega,\mathcal{F},\mathbb{P})$ and a lattice property of the set of trading strategies) there exists an optimal strategy $\tilde{\xi}_T(x)$ such that

$$U_{T-1}(x) = \mathbf{E}\left[U_T\left(x + \left\langle\tilde{\xi}_T(x), \Delta S_T\right\rangle\right)|\mathcal{F}_{T-1}\right]. \tag{2.4.30}$$

The method to prove that $\tilde{\xi}_T$ must satisfy the inequality (2.4.29) is rather indirect. Following the proof of Lemma 4.8 in [44] we claim that there is a r.v. K_{T-1} such that $|\xi| \geqslant K_{T-1}$ implies

[8] See Appendix B, Lemma B.2.1

that

$$\mathbf{E}\left[U\left(x+\langle\xi,\Delta S_T\rangle\right)\middle|\mathcal{F}_{T-1}\right]\leqslant U\left(x\right), \tag{2.4.31}$$

therefore $\xi=0$ would be a better trading strategy. Moreover, we can substitute a r.v. $\xi\in\Xi_{T-1}$ by $\tilde{\xi}$ its orthogonal projection into the subspace $D_T\left(\omega\right)$.

Lemma 2.4.18. *Suppose* $V:\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ *an* $\mathcal{F}_{T-1}\otimes\mathcal{B}\left(\mathbb{R}^d\right)-$*measurable. Let* $\xi\in\Xi_{T-1}$, *denote by* $\hat{\xi}$ *its orthogonal projection into* $D_T\left(\omega\right)$. *Then*

$$\mathbf{E}\left[V\left(x+\langle\xi,\Delta S_T\rangle\right)\middle|\mathcal{F}_{T-1}\right]=\mathbf{E}\left[V\left(x+\left\langle\hat{\xi},\Delta S_T\right\rangle\right)\middle|\mathcal{F}_{T-1}\right], \tag{2.4.32}$$

for each $x\in\mathbb{R}$.

We do not include the proof of this fact in this chapter, but it is given in the appendix C, Proposition 2.4.18, as the proof depends on results that are relevant only in the proof of the lemma and not applied anywhere else in the subsequent part of the chapter.

Thus, hereafter in this section we shall consider trading strategies in $\xi\in\Xi_{T-1}$ such that $\xi\in D_T\left(\omega\right)$.

Proposition 2.4.19. *Let* $x\in\mathbb{R}$. *Suppose all the assumptions in Theorem 2.4.16 hold, then there is a random variable* $K_{T-1}\left(x,\omega\right)$ *such that, if* $\xi_T\left(\omega\right)$ *is a trading strategy with* $|\xi_T|\geqslant K_{T-1}\left(x,\omega\right)$ *then*

$$\mathbf{E}\left[U_T\left(x+\langle\xi_T,\Delta S_T\rangle\right)\middle|\mathcal{F}_{T-1}\right]\leqslant\mathbf{E}\left[U_T\left(x\right)\middle|\mathcal{F}_{T-1}\right]. \tag{2.4.33}$$

Moreover, $K_{T-1}\left(x,\omega\right)$ *can be taken of the form* $\psi_{T-1}\left(1+|x|\right)^{\zeta_{T-1}}$ *with* $\psi_{T-1}\in\mathscr{W}$.

We shall prove the claim in these steps, and only for '$t+1$' $=T$, for a complete proof we apply an inductive argument. Notice that, by the definition of U_T (see definitions in Assumption 2.4.2, (2.4.17) and (**??**)) we can interchange the $U_T\left(x\right)$ by $U\left(x\right)$

(i) Bound the function $U_T^-\left(x+\langle\xi,\Delta S_T\rangle\right)$ from below if $\xi\in D_T\left(\omega\right)$ on a set that depends on U_T;

(ii) Bound the function $U_T^+\left(x+\langle\xi,\Delta S_T\rangle\right)$ from above by an expression of 'order' $|\xi|^\gamma$ for $\xi\in\Xi_{T-1}$ $|\xi|\geqslant K_{T-1}\left(\omega\right)$;

(iii) show that the condition on (i) can be satisfied if $|\xi|\geqslant\psi_{T-1}\left(1+|x|^{\alpha_{T-1}}\right)$ and $\psi_{T-1}\in\mathscr{W}$.

Finally the bounds in (i), (ii) and (iii) allow to obtain 2.4.33. We shall check the steps $(i)-(iii)$ for $t=T-1$, explaining the arguments in [44] in more detail. The proofs in the general case $t\in\{1,\ldots,T-1\}$ follow by induction.

Lemma 2.4.20. *Given that Assumptions 2.4.1, 2.4.2 and those in Theorem 2.4.16 are in force, there exists a measurable set* B_T *such that if,* $\xi_T\in\Xi_{T-1}$, $|\xi_T|\geqslant 1$ *such that*

$$\mathbf{E}\left[U^-\left(x+\langle\xi,\Delta S_T\rangle\right)\middle|\mathcal{F}_{T-1}\right]\geqslant|\xi|^{\frac{1+\gamma}{2}}\,\mathbb{P}\left(B_T\middle|\mathcal{F}_{T-1}\right)-c\,|\xi|^\gamma, \tag{2.4.34}$$

Proof. Define

$$B_T:=\left\{\left\langle\frac{\xi}{|\xi|},\Delta S_T\right\rangle<-\beta_{T-1},\,U\left(\frac{x}{|\xi|^{\frac{1+\gamma}{2}}}-|\xi|^{\frac{1-\gamma}{2}}\beta_{T-1}\right)<-1\right\},$$

By Proposition 2.4.12 we have

$$c\lambda^{\gamma} - U(\lambda z) \geqslant \lambda(-U(z)),$$

for $\lambda \geqslant 1$ applying this to $\lambda = |\xi|^{\frac{1+\gamma}{2}}$ and to $z = \frac{x}{|\xi|^{\frac{1+\gamma}{2}}} + \left\langle \frac{\xi}{|\xi|}, \Delta S_T \right\rangle |\xi|^{\frac{1-\gamma}{2}}$,

$$-|\xi|^{\frac{1+\gamma}{2}} U\left(\frac{x}{|\xi|^{\frac{1+\gamma}{2}}} + \left\langle \frac{\xi}{|\xi|}, \Delta S_T \right\rangle |\xi|^{\frac{1-\gamma}{2}}\right) \leqslant c\,|\xi|^{\gamma\frac{1+\gamma}{2}} - U\left(\frac{x}{|\xi|^{(1+\gamma)/2}} + \left\langle \frac{\xi}{|\xi|}, \Delta S_T \right\rangle |\xi|\right). \tag{2.4.35}$$

As $0 < \gamma < 1$ then we can bound from above the term in (2.4.35) $|\xi|^{\gamma\frac{1+\gamma}{2}}$ by $|\xi|^{\gamma}$. And then $\mathbb{I}_{B_T}\Big(c\,|\xi|^{\gamma} + U(x + \langle \xi, \Delta S_T \rangle)\Big) \geqslant |\xi|^{\frac{1+\gamma}{2}} \mathbb{I}_{B_T}$, note that on B_T, we have $U\left(x + \left\langle \frac{\xi}{|\xi|}, \Delta S_T \right\rangle\right) < U(-\beta_{T-1}) \leqslant 0$ and $c\,|\xi|^{\gamma} - U(x + \langle \xi, \Delta S_T \rangle) > 0$, thus

$$U^{-}(x + \langle \xi, \Delta S_T \rangle) \geqslant |\xi|^{\frac{1+\gamma}{2}} \mathbb{I}_{B_T} - c\,|\xi|^{\gamma}, \tag{2.4.36}$$

and as $\xi_T \in \Xi_{T-1}$

$$-\mathbf{E}\left[U^{-}(x + \langle \xi, \Delta S_T \rangle) \,\middle|\, \mathcal{F}_{T-1}\right] \leqslant c\,|\xi|^{\gamma} - |\xi|^{\frac{1+\gamma}{2}} \mathbb{P}(B_T \,|\, \mathcal{F}_{T-1}). \tag{2.4.37}$$

□

For ease of reference we denote A_T the set

$$A_T = \left\{\omega \in \Omega \,:\, \operatorname*{ess\,inf}_{q \in \tilde{\Xi}_{T-1}} \mathbb{P}\left(U\left(x - |\xi|^{\frac{1-\gamma}{2}} \beta_{T-1}\right) < -1,\ \langle q, \Delta S_T \rangle < -\beta_{T-1} \middle| \mathcal{F}_{T-1}\right) \geqslant \frac{1}{2}\kappa_{T-1}\right\}.$$

If

$$\operatorname*{ess\,inf}_{q \in \tilde{\Xi}_{T-1}} \mathbb{P}\left(\langle q, \Delta S_T \rangle < -\beta_{T-1}\,, U\left(\frac{x}{|\xi|^{\frac{1+\gamma}{2}}} - \beta_{T-1}\,|\xi|^{\frac{1-\gamma}{2}}\right) < -1 \middle| \mathcal{F}_T\right) \geqslant \frac{1}{2}\kappa_{T-1}. \tag{2.4.38}$$

Then, on A_T we have

$$-\mathbf{E}\left[U^{-}(x + \langle \xi, \Delta S_T \rangle) \,\middle|\, \mathcal{F}_{T-1}\right] \leqslant c\,|\xi|^{\frac{1+\gamma}{2}} - \frac{1}{2}\,|\xi|^{\frac{1+\gamma}{2}} \kappa_{T-1}. \tag{2.4.39}$$

Now we turn our attention to *ii*).

Lemma 2.4.21. *Let $\xi_T \in \Xi_{T-1}$ and suppose that the Assumptions 2.4.1, 2.4.2 and the assumptions in Theorem 2.4.16 hold, then there is a random variable $G_T(\omega) \in \mathscr{W}$ such that*

$$\mathbf{E}\left[U^{+}(x + \langle \xi_T, \Delta S_T \rangle) \,\middle|\, \mathcal{F}_{T-1}\right] \leqslant G_T(\omega)\,|\xi|^{\gamma}.$$

Proof. By Assumption 2.4.1 and (2.4.12) and (2.4.13) for $\xi \in \Xi_{T-1}$ with $\lambda = |\xi| \geqslant 1$,

$$U^{+}(x + \langle \xi, \Delta S_T \rangle) = U^{+}\left(|\xi|\left(x/|\xi| + \left\langle \frac{\xi}{|\xi|}, \Delta S_T \right\rangle\right)\right) \leqslant |\xi|^{\gamma} U^{+}\left(x + \left\langle \frac{\xi}{|\xi|}, \Delta S_T \right\rangle\right) + C\,|\xi|^{\gamma}, \tag{2.4.40}$$

and the last term can be bounded by a measurable finite function. Indeed, define $\mathcal{S} = \left\{\theta \in \mathbb{R}^d : \theta_k \in \{1, -1\} \text{ for all } k = 1, \ldots, d\right\}$. As U^{+} is non-decreasing, we can get the following estimate

$$\mathbf{E}\left[U^{+}\left(x+\left\langle\frac{\xi}{|\xi|},\Delta S_T\right\rangle\right)\Big|\mathcal{F}_{T-1}\right]\leqslant\mathbf{E}\left[\max_{\theta\in\mathcal{S}}U^{+}\left(x+\langle\theta,\Delta S_T\rangle\right)\Big|\mathcal{F}_{T-1}\right],$$

and $\mathbf{E}\left[\max_{\theta\in\mathcal{S}}U^{+}\left(x+\langle\theta,\Delta S_T\rangle\right)\Big|\mathcal{F}_{T-1}\right]\leqslant\sum_{\theta\in\mathcal{S}}\mathbf{E}\left[U^{+}\left(x+\langle\theta,\Delta S_T\rangle\right)\Big|\mathcal{F}_{T-1}\right]$.
Define

$$L_T(\omega):=\sum_{\theta\in\mathcal{S}}\mathbf{E}\left[U^{+}\left(x+\langle\theta,\Delta S_T\rangle\right)|\mathcal{F}_{T-1}\right]. \tag{2.4.41}$$

We claim that $L_T(\omega)$ is integrable.
Indeed, as $\Delta S_T\in\mathscr{W}$ by Cauchy-Schwarz inequality and our assumptions on U

$$\mathbf{E}\left[U^{+}\Big(x+\langle\theta,\Delta S_T\rangle\Big)\big|\mathcal{F}_{T-1}\right]\leqslant\mathbf{E}\left[U^{+}\Big(|x|+\sqrt{d}\,|\Delta S_T|\Big)\big|\mathcal{F}_{T-1}\right], \tag{2.4.42}$$

$$L_T(\omega)\leqslant\sum_{\theta\in\mathcal{S}}\mathbf{E}\left[U^{+}\big(|x|+|\theta|\,|\Delta S_T|\big)\big|\mathcal{F}_{T-1}\right]=2^d\mathbf{E}\left[U^{+}\left(|x|+\sqrt{d}\,|\Delta S_T|\right)|\mathcal{F}_{T-1}\right].$$

Recall that, by assumption, $U(x)$ is continuously differentiable, strictly increasing and $U(0)=0$, then $U(\varepsilon)>0$ for any positive number $\varepsilon>0$. Fixing ω and applying the mean value theorem on $\left[\epsilon,|x|+\sqrt{d}\,|\Delta S_T|\right]$ and then letting $\epsilon\downarrow 0$ for some $\epsilon\leqslant\varsigma\leqslant|x|+\sqrt{d}\,|\Delta S_T|$

$$U^{+}\left(|x|+\sqrt{d}\,|\Delta S_T|\right)-U^{+}(\epsilon)=U'(\varsigma)\left(|x|+\sqrt{d}\,|\Delta S_T|-\epsilon\right)\leqslant U'(\epsilon)\left(|x|+\sqrt{d}\,|\Delta S_T|\right),$$

$$L_T(\omega)\leqslant 2^dU'(0)\,\mathbf{E}\left[|x|+\sqrt{d}\,|\Delta S_T|\,|\mathcal{F}_{T-1}\right]\leqslant 2^dU'(0)\left(|x|+\sqrt{d}\mathbf{E}\left[|\Delta S_T|\,|\mathcal{F}_{T-1}\right]\right),$$

as the last term in the last inequality belongs to $\mathscr{W}$ then $L_T\in\mathscr{W}$.
Furthermore, by (2.4.40) and the previous estimation given above

$$\mathbf{E}\left[U^{+}\left(x+\langle\xi,\Delta S_T\rangle\right)|\mathcal{F}_{T-1}\right]\leqslant L_T(\omega)\,|\xi|^{\gamma}+C\,|\xi|^{\gamma}. \tag{2.4.43}$$

□

In the case $|\xi|\geqslant 1$, and $\omega\in A_T$ we can bound the 'expected future wealth' by a random power function of $|\xi|$

$$\mathbf{E}\left[U\big(x+\langle\xi,\Delta S_T\rangle\big)\big|\mathcal{F}_{T-1}\right]\leqslant L_T(\omega)\,|\xi|^{\gamma}+2C\,|\xi|^{\gamma}-\frac{1}{2}\,|\xi|^{\frac{1+\gamma}{2}}\,\kappa_{T-1}. \tag{2.4.44}$$

Choosing ξ such that $|\xi|>M(x,\omega)$ (specified below) it is possible to obtain

$$L_T(\omega)\,|\xi|^{\gamma}+2C\,|\xi|^{\gamma}-\frac{1}{2}\,|\xi|^{\frac{1+\gamma}{2}}\,\kappa_{T-1}\leqslant\mathbf{E}\left[U_T(x)\,|\mathcal{F}_{T-1}\right].$$

In this case, the last term is equal to $U(x)$ and it changes accordingly in the case $t\leqslant T-1$.
Now we proceed with the step (iii), we see that there is a r.v. $K'(\omega,x)\in\mathscr{W}$ such that, if $|\xi(\omega)|\geqslant K'(\omega,x)$ (we shall prove that $K'(\omega,x)=\psi_T\left(1+|x|^{\alpha_T}\right)$ with $\psi_T\in\mathscr{W}$ is a possible choice) then

$$\operatorname*{ess\,inf}_{q\in\Xi_{T-1}}\mathbb{P}\left(\langle q,\Delta S_T\rangle<-\beta_{T-1},\,U\left(\frac{x}{|\xi|^{\frac{1+\gamma}{2}}}-\beta_{T-1}\,|\xi|^{\frac{1+\gamma}{2}}\right)<-1\middle|\mathcal{F}_{T-1}\right)\geqslant\frac{1}{2}\kappa_{T-1}. \tag{2.4.45}$$

We claim that if $|\xi|$ is 'large' enough then

$$\operatorname*{ess\,inf}_{q\in\tilde{\Xi}_{T-1}} \mathbb{P}\left(\langle q,\Delta S_T\rangle < -\beta_{T-1},\ U\left(\frac{x}{|\xi|^{(1+\gamma)/2}} - \beta_{T-1}|\xi|^{\frac{1-\gamma}{2}}\right) < -1\middle|\mathcal{F}_{T-1}\right) \geqslant \frac{1}{2}\kappa_{T-1}\ \ \mathbb{P}-a.s. \tag{2.4.46}$$

First of all, if $x \geqslant 0$ then $U(x) \geqslant 0$, thus let us take $|\xi|^{\frac{1-\gamma}{2}}\beta_{T-1} \geqslant |x|$ *a.s.* by concavity

$$U\left(|x| - |\xi|^{\frac{1-\gamma}{2}}\beta_{T-1}\right) \leqslant U(0) - U'(0)\left(|\xi|^{\frac{1-\gamma}{2}}\beta_{T-1} - |x|\right). \tag{2.4.47}$$

Let ψ_T and ρ_T be positive and belonging to $\mathscr{W}$, as U' is non-increasing then $U'(0) \geqslant \mathbf{E}\left[U'(\psi_T)|\mathcal{F}_{T-1}\right]$ and

$$U\left(|x| - |\xi|^{\frac{1-\gamma}{2}}\beta_{T-1}\right) \leqslant \mathbf{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] - \mathbf{E}\left[U'(\psi_T)|\mathcal{F}_{T-1}\right]\left(|\xi|^{\frac{1-\gamma}{2}}\beta_{T-1} - |x|\right). \tag{2.4.48}$$

Denote by $N_1^{(T-1)} := 4\mathbf{E}\left[\psi_T|\mathcal{F}_{T-1}\right]/\kappa_{T-1}$ and $N_2^{(T-1)} := 4\mathbf{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right]/\kappa_{T-1}$. By the theorem on existence of regular conditional probabilities see Appendix Theorem A.2.2. There is a version of $\mathbb{P}(\cdot|\mathcal{F}_{T-1})(\omega)$ denoted by $Q_{T-1}(\omega,\cdot)$ such that, except for a null subset $N \in \mathcal{F}$, for any $\omega \in \Omega/N$ the mapping $Q_{T-1}(\omega,\cdot)$ is a probability measure on $\mathcal{F}_{T-1}$,

$$\begin{aligned} Q\Big(\omega : \left\{\psi_T > N_1^{(T-1)}\right\} \cup \left\{\mathbb{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] > N_2^{(T-1)}\right\}\Big) \leqslant\, & Q\left(\omega : \psi_T > N_1^{(T-1)}\right) \\ & + Q\Big(\omega : \mathbb{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] > N_2^{(T-1)}\Big), \end{aligned} \tag{2.4.49}$$

applying Markov's inequality to the sum on the right hand side of the inequality

$$Q\left(\omega : \psi_T > N_1^{(T-1)}\right) \leqslant \frac{1}{N_1^{(T-1)}}\mathbb{E}\left[\psi_T|\mathcal{F}_{T-1}\right] \leqslant \frac{1}{4}\kappa_{T-1}, \tag{2.4.50}$$

$$Q\left(\omega : \mathbb{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] > N_2^{(T-1)}\right) \leqslant \frac{1}{N_2^{(T-1)}}\mathbb{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] \leqslant \frac{1}{4}\kappa_{T-1}.$$

Then

$$\mathbb{P}\left(\psi_T \leqslant N_1^{(T-1)},\ \mathbb{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] \leqslant N_2^{(T-1)}\middle|\mathcal{F}_{T-1}\right) \geqslant 1 - \frac{1}{2}\kappa_{T-1}, \tag{2.4.51}$$

by the estimates in (2.4.47) and (2.4.48)

$$\left\{\mathbb{E}\left[U'(\psi_T)|\mathcal{F}_{T-1}\right]\left(|\xi|^{\frac{1-\gamma}{2}}\beta_{T-1} - |x|\right) - \mathbf{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] > 1\right\} \subset \left\{-U\left(|x| - |\xi|^{\frac{1-\gamma}{2}\beta_{T-1}}\right) > 1\right\}.$$

Define the set Γ_{T-1}

$$\begin{aligned} \Gamma_{T-1} := \Big\{U'\left(N_1^{(T-1)}\right)\left(\beta_{T-1}|\xi|^{\frac{1-\gamma}{2}} - |x|\right) - N_2^{(T-1)} > 1,\ \mathbf{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] \leqslant N_2^{(T-1)}, \\ \psi_T \leqslant N_1^{(T-1)},\ \langle p,\Delta S_T\rangle < -\beta_{T-1}\Big\}. \end{aligned}$$

On Γ_{T-1} we have the following inequalities

$$1 < \mathbf{E}\left[U'(\psi_T)|\mathcal{F}_{T-1}\right]\left(\beta_{T-1}|\xi|^{\frac{1-\gamma}{2}} - |x|\right) - \mathbf{E}\left[U(\rho_T)|\mathcal{F}_{T-1}\right] \leqslant -U\left(|x| - |\xi|^{\frac{1-\gamma}{2}}\beta_{T-1}\right).$$

In addition to the last inequalities, previous estimates give

$$\mathbb{P}\left(\Gamma_{T-1}\left|\mathcal{F}_{T-1}\right.\right) \geqslant \mathbb{P}\left(\psi_T \leqslant N_1^{(T-1)},\ \mathbf{E}\left[U\left(\rho_T\right)\left|\mathcal{F}_{T-1}\right.\right] \leqslant N_2^{(T-1)}\left|\mathcal{F}_{T-1}\right.\right) + \mathbb{P}\left[U'\left(N_1^{(T-1)}\right)\left(\beta_{T-1}\left|\xi\right|^{\frac{1-\gamma}{2}} - |x|\right) > 1 + N_2^{(T-1)},\ \langle q, \Delta S_T\rangle < -\beta_{T-1}\left|\mathcal{F}_{T-1}\right.\right]. \tag{2.4.52}$$

If the first event on the second term in (2.4.52) is 'sure to happen', we can readily apply the estimate obtained in Theorem 2.4.9, this justifies that if the trading strategy is such that for all $\omega \in \Omega$

$$|\xi| > \left[\left(\frac{1 + N_2^{(T-1)}}{U'\left(N_1^{(T-1)}\right)} + |x|\right)\frac{1}{\beta_{T-1}}\right]^{\frac{2}{1-\gamma}}, \tag{2.4.53}$$

as $N_1^{(T-1)},\ N_2^{(T-1)}$, ξ are $\mathcal{F}_{T-1}-$measurable. Then

$$\begin{aligned}\operatorname*{ess\,inf}_{q\in\tilde{\Xi}_{T-1}} \mathbf{P}\left(U'\left(N_1^{(T-1)}\right)\left(\beta_{T-1}\left|\xi\right|^{\frac{1-\gamma}{2}} - |x|\right) > 1 + N_2^{(T-1)},\ \langle q, \Delta S_T\rangle < -\beta_{T-1}\left|\mathcal{F}_{T-1}\right.\right)\\ = \operatorname*{ess\,inf}_{q\in\tilde{\Xi}_{T-1}} \mathbb{P}\left(\langle q, \Delta S_T\rangle < -\beta_{T-1}\left|\mathcal{F}_{T-1}\right.\right) \geqslant \kappa_{T-1},\end{aligned}$$

it follows by (2.4.51) and (2.4.52) that

$$\mathbb{P}\left(\Gamma_{T-1}\left|\mathcal{F}_{T-1}\right.\right) \geqslant 1 - \frac{1}{2}\kappa_{T-1} + \kappa_{T-1} - 1 = \frac{1}{2}\kappa_{T-1}.$$

Notice that it is enough to have $|\xi| > \psi_T\left(1 + |x|^{\varsigma_T}\right)$ where $\varsigma_T = \frac{2}{1-\gamma}$ and

$$J = \left(\left(1 + N_2^{(T-1)}\right)\left|N_1^{(T-1)}\right|^l \cdot \frac{1}{M} \vee 1\right)^{\varsigma_T} \frac{1}{\beta_{T-1}^{\varsigma_T}}.$$

Finally, define

$$\psi_T(\omega) := J \vee \left[\frac{12C}{\kappa_{T-1}} \vee \frac{6L_T}{\kappa_{T-1}}\right]^{\frac{2}{1-\gamma}} \vee \left(\frac{6}{\kappa_{T-1}} \cdot \mathbf{E}\left[U\left(-x^-\right)\middle|\mathcal{F}_{T-1}\right]^-\right)^{\frac{2}{1+\gamma}}, \tag{2.4.54}$$

for any $\xi \in \Xi_{T-1}$ such that $|\xi| \geqslant \psi_T(\omega)\left(1 + |x|^{\varsigma_T}\right)$ it follows that

$$\mathbf{E}\left[U\left(x + \langle\xi, \Delta S_T\rangle\right)\middle|\mathcal{F}_{T-1}\right] \leqslant L_T(\omega)\left|\xi\right|^\gamma + 2C\left|\xi\right|^\gamma - \frac{1}{2}\left|\xi\right|^{\frac{1+\gamma}{2}}\kappa_{T-1} \leqslant \mathbf{E}\left[U(x)\middle|\mathcal{F}_{T-1}\right].$$

This implies that if $\xi \in D_T(\omega)$ is optimal then we must have $J_T, K' \in \mathscr{W}$ and a number $\varsigma_T := \frac{2}{1-\gamma} \vee 1 > 0$ such that

$$|\xi| \leqslant \psi_T\left(1 + |x|^{\varsigma_T}\right). \tag{2.4.55}$$

This is the claim in [44] section 7, Proposition 7.1, (42).

Lemma 2.4.22. *Under Assumptions 2.4.2, 2.4.1 and the assumptions of Theorem 2.4.16. The function $U_{T-1}(x)$ is finite $\mathbb{P}-a.s.$*

Proof. As U is increasing and $U(0) = 0$ and (2.4.23) by assumption in Theorem 2.4.16.

$$\begin{aligned}U\left(x + |y|\left|\Delta S_T\right|\right) &\leqslant \int_0^{|x|+|y||\Delta S_T|} K\left(t^k + 1\right)dt, \\ &\leqslant K|x| + K|y|\left|\Delta S_T\right| + \left(2^{k-1} \vee 1\right)\frac{K}{k+1}\left(|x|^{k+1} + |y|^{k+1}\left|\Delta S_T\right|^{k+1}\right),\end{aligned} \tag{2.4.56}$$

if $\tilde{\xi}$ satisfies (2.4.55), the last estimate (2.4.56) applied to $y = \psi_T\left(1 + |x|^{\varsigma_T}\right)$ implies

$$\mathbf{E}\left[U\left(x + \left\langle \tilde{\xi}, \Delta S_T \right\rangle\right) |\mathcal{F}_{T-1}\right] \leqslant \mathbf{E}\left[U\Big(|x| + \psi_T |\Delta S_T| \left(1 + |x|^{\varsigma_T}\right)\Big) |\mathcal{F}_{T-1}\right]. \tag{2.4.57}$$

We obtain that $\mathbf{E}\left[U\left(x + \left\langle \tilde{\xi}, \Delta S_T \right\rangle\right) \middle| \mathcal{F}_{T-1}\right]$ is bounded by a sum of powers of random variables J_T, $|\Delta S_T| \in \mathscr{W}$. On the other hand there exists a random variable $\tilde{\xi}_T \in \Xi_{t-1}$ such that (2.4.30) and condition (2.4.55) hold. Thus, it follows that $U_{T-1}(x) < \infty$ $\mathbb{P} - a.s.$.

□

As U is continuously differentiable, the following lemma states that regularity is preserved in dynamic programming, i.e. the function $U_{T-1}(x)$ is (for fixed ω) continuously differentiable.

Lemma 2.4.23. *Under the assumptions in Theorem 2.4.16, the functions $U_t(x)$ are continuously differentiable for any $t \in \{1, 2, \ldots, T\}$ i.e. a version of $U'_t\big(x, \omega\big)$ is given by* $\operatorname{ess\,sup}_{\xi \in \Xi_t} \mathbf{E}\big[U'_{t+1}\big(x + \langle \xi, \Delta S_t \rangle\big) \big| \mathcal{F}_t\big]$

Proof. The existence of a r.v. $\tilde{\xi}_T(x) \in \Xi_{T-1}$ such that (2.4.30) holds and the last estimates in (2.4.57) imply

$$U_{T-1}(x) \leqslant U'(0)\, \mathbf{E}\left[|x| + \rho_{T-1}\left(1 + |x|^{\varsigma_{T-1}}\right) |\mathcal{F}_{T-1}\right] \in \mathscr{W}, \tag{2.4.58}$$

This shall follow from Proposition 6.4 in [44], but in order to apply this result we must ensure that $-\infty < \mathbf{E}\left[U\left(x + \langle y, \Delta S_T \rangle\right)\right] < \infty$ and

$$\sup_{x, |y| \in [a,b]} U'\big(x + |y|\, |\Delta S_T|\big) |\Delta S_T| \in L^1. \tag{2.4.59}$$

The first condition follows from (2.4.56) and (2.4.59). Indeed, (2.4.56) shows in particular that $\mathbf{E}U\left(x + \langle y, \Delta S_T \rangle\right) < \infty$. On the other hand, we have by (2.4.23) and the fact that U is concave,

$$\mathbf{E}U\left(x + \langle y, \Delta S_T \rangle\right) \geqslant -\mathbf{E}U\left(-|x| - |y|\, |\Delta S_T|\right) = \mathbf{E}\int_{-|x|-|y||\Delta S_T|}^{-|x|} U'(s)\, ds - U\left(-|x|\right),$$

as $U(-|x|) < \infty$ it is enough to see that the r.v. involving the integral of U' has finite, but as U is concave, U' is decreasing,

$$\mathbf{E}\int_{-|x|-|y||\Delta S_T|}^{-|x|} U'(s)\, ds \geqslant \mathbf{E}\big(U'\left(-|x|\right)\big) |y|\, |\Delta S_T|,$$

then (2.4.59) implies $\mathbf{E}U\left(x + \langle y, \Delta S_T \rangle\right) > -\infty$. In the case $t = T-1$, (2.4.59) is a consequence of Assumption 2.4.23 and assumptions in Theorem 2.4.16.
The assumption on ΔS_T in Proposition 6.4 in [44], can be replaced by $\Delta S_T \in \mathscr{W}$ without changes in its proof. □

For the general case i.e. $U'_t(x)$, Proposition 6.4 and 6.5 in [44] leads to

$$U'_t(x) = \mathbf{E}\left[U'_{t+1}\left(x + \left\langle \tilde{\xi}_t(x), \Delta S_{t+1} \right\rangle\right) |\mathcal{F}_t\right], \tag{2.4.60}$$

if the following condition holds

$$\sup_{x, |y| \in [a,b]} U'_{t+1}\big(x + |y|\, |\Delta S_{t+1}|\big) |\Delta S_{t+1}| \in L^1. \tag{2.4.61}$$

An inductive argument allows to see that, assuming that (2.4.60) holds for U'_{t+1}, there are $\tilde{\xi}(x)_i$, $\psi_i \in \mathscr{W}$ and constants ς_i for $i \geqslant t+1$ satisfying (2.4.55). Using these variables, one can verify that there is a polynomial $p_{t+1}(w,z)$ such that

$$U'_{t+1}\left(x+\langle y,\Delta S_{t+1}\rangle\right)|\Delta S_{t+1}| \leqslant |\Delta S_{t+1}|\, U'\big(-p_{t+1}\left(\chi_{t+1},\Delta S_{t+1}\right)\big), \tag{2.4.62}$$

and $\chi_{t+1} \in \mathscr{W}$. By assumption on U' we can bound the last part in (2.4.62) by an element belonging to $\mathscr{W}$ (as any polynomial of random variables in $\mathscr{W}$ belongs to $\mathscr{W}$). We explain the argument for when $t = T-1$ but the same reasoning applies when $t < T-1$. Applying Proposition 6.4 in [44] we obtained that $U_{T-1}(x)$ is continuously differentiable $\mathbb{P}-a.s.$, and $\mathbf{E}\left[U_T\left(x+\langle y,\Delta S_T\rangle\right)\middle|\mathcal{F}_{T-1}\right]$ is continuously differentiable in $(x,y) \in \mathbb{R}^{d+1}$, as well. Furthermore, by eq. (37) in Proposition 6.5 and Proposition 6.6 in [44] there is $\tilde{\xi}_T(x) \in \Xi_{T-1}$ such that we can express $U'_{T-1}(x)$ as

$$U'_{T-1}(x) = \mathbf{E}\left[U'_T\left(x+\left\langle\tilde{\xi}_T(x),\Delta S_T\right\rangle\right)\middle|\mathcal{F}_{T-1}\right]. \tag{2.4.63}$$

In other words, a version of $U'_{T-1}(x)$ is given by $\mathbf{E}\left[U'_T\left(x+\left\langle\tilde{\xi}_T(x),\Delta S_T\right\rangle\right)\middle|\mathcal{F}_{T-1}\right]$.

$$U_T\left(x+\left\langle\tilde{\xi}_T(x),\Delta S_T\right\rangle\right)|\Delta S_T| \leqslant U(|x|) + K\left[\left(|x|+\left|\tilde{\xi}_T(x)\right||\Delta S_T|\right)^k+1\right]|\Delta S_T|,$$

and by the condition (2.4.55)

$$K\left[\left(|x|+\left|\tilde{\xi}_T(x)\right||\Delta S_T|\right)^k+1\right]|\Delta S_T| \leqslant K\left[\left(|x|+\psi_T\left(1+|x|^{\varsigma_T}\right)|\Delta S_T|\right)^k+1\right]|\Delta S_T|.$$

And $\left(2^{k-1}\vee 1\right)K\left(\psi_T^k\left(1+|x|^{\varsigma_T k}\right)\right)|\Delta S_T|^k \in \mathscr{W}$.

For fixed ω the function $U_{T-1}(x)$ is continuously differentiable by an application of Proposition 6.4. Applying the arguments given above, we can prove that $U_t(x)$ is continuously differentiable $\mathbb{P}-a.s.$
We have sketched the main arguments from [44], showing that the 'value functions' $U_t(x)$ satisfy the following properties:

- For some $\rho_t \in \mathscr{W}$ and a constant ς_t,

$$U_t(x) \leqslant \mathbf{E}\left[U_{t+1}\left(x+\left(1+|x|^{\varsigma_t}\right)\rho_t\right)\middle|\mathcal{F}_t\right]. \tag{2.4.64}$$

- $U_t(x) < \infty$ for all $x \in \mathbb{R}$, $\mathbb{P}-a.s.$

- There is a trading strategy (optimal) $\tilde{\xi}_{t+1}(x) \in \Xi_t$ and $\tilde{\xi}_{t+1} \in D_t(\omega)$ such that

$$U_t(x) = \mathbf{E}\left[U_{t+1}\left(x+\left\langle\tilde{\xi}_{t+1}(x),\Delta S_{t+1}\right\rangle\right)\middle|\mathcal{F}_t\right]. \tag{2.4.65}$$

  and

$$\left|\tilde{\xi}(x)_{t+1}(x)\right| \leqslant \psi_t\left(1+|x|^{\varsigma_t}\right).$$

- The function $U_t(x)$ is continuously differentiable and

$$U'_t(x) = \mathbf{E}\left[U'_{t+1}\left(x+\left\langle\tilde{\xi}_{t+1}(x),\Delta S_{t+1}\right\rangle\right)\middle|\mathcal{F}_t\right].$$

From (2.4.65) we can obtain (by an approximation of a random variables by simple ones) we can prove that if X is a random variable, then

$$U_t(X) = \mathbf{E}\left[U\left(X + \left\langle\tilde{\xi}_{t+1}(X), \Delta S_{t+1}\right\rangle\right)|\mathcal{F}_t\right]. \tag{2.4.66}$$

4.4 Dynamic Programing and martingale measures

We include the following proposition whose proof can be found in [44], we follow their arguments and elaborate further.

Proposition 2.4.24. *The functions $U_t(x)$, $t \in \{0, 1, \ldots, T\}$, have continuously differentiable versions such that $-\infty < \mathbf{E}U_t(x + \langle y, \Delta S_{t+1}\rangle) < \infty$. Furthermore, for $1 \leqslant i \leqslant d$ and for $1 \leqslant t \leqslant T$*

$$\mathbf{E}\left[U_t'\left(X + \left\langle\tilde{\xi}_t(X), \Delta S_t\right\rangle\right)\Delta S_t^i|\mathcal{F}_{t-1}\right] = 0, \tag{2.4.67}$$

for any $\mathcal{F}_{t-1}-$measurable random variable X.

As it was previously mentioned, Proposition 5.2 in [44] allows to express $U_t(x, \omega)$ as the maximum utility on a one period investment using $\tilde{\xi}_{t+1}(x)$, and this strategy is dominated by a r.v. in $\mathscr{W}$, see (2.4.55).
Proposition 2.4.24 implies that by *using the 'one-step' trading strategy* $\tilde{\xi}_{t+1}(x)$, the equality of versions in (2.4.67) holds. This identity resembles the 'critical point condition' in elementary calculus to obtain a local maximum/minimum.
Moreover, define recursively the trading strategy $\phi^\star$ given by $\phi^\star(x) := \left\{\tilde{\xi}_i\left(V_{i-1}^{x,\phi}\right)\right\}_{1\leqslant i\leqslant T}$ with $V_0^{x,\phi^\star} = x$. For instance, $\phi_1^\star(x) = \tilde{\xi}_1(x)$, where $\tilde{\xi}_1(x)$ is the strategy that enables equality in (2.4.65), this generates a wealth given by $V_1^{x,\phi_1^\star} = x + \left\langle\tilde{\xi}_1(x), \Delta S_1\right\rangle$. Then, define $\phi_2^\star(x) := \tilde{\xi}_2\left(V_1^{x,\phi_1^\star}\right)$ choose the trading stategy in (2.4.66) (and $X = V_1^{x,\phi_1^\star}$) again this produces a wealth equal to $V_2^{x,\phi_2^\star} = V_1^{x,\phi_1^\star} + \left\langle\tilde{\xi}_2\left(V_1^{x,\phi_1^\star}\right), \Delta S_2\right\rangle$, notice that indeed, $V_2^{x,\phi_2^\star}$ is the portfolio value process of the trading strategy $\phi_2^\star(x) = \left(\tilde{\xi}_1(x), \tilde{\xi}_2\left(V_1^{x,\phi_1^\star}\right)\right)$ and define $\phi_3^\star(x) = \tilde{\xi}_3\left(V_2^{x,\phi_2^\star}\right)$ iterating this procedure, yields the trading strategy $\{\phi_t^\star\}_{t\geqslant 1}$

$$\begin{aligned}\mathbf{E}\left[U'\left(V_T^{x,\phi^\star}\right)|\mathcal{F}_{T-1}\right] &= \mathbf{E}\left[U'\big(x + \sum_{i=1}^{T-1}\langle\phi_i^\star(x), \Delta S_i\rangle + \langle\phi_T^\star(x), \Delta S_T\rangle\big)\big|\mathcal{F}_{T-1}\right] = \\ &= U_{T-1}'\big(x + \sum_{i=1}^{T-1}\langle\phi_i^\star, \Delta S_i\rangle\big),\end{aligned} \tag{2.4.68}$$

by (2.4.60) and taking conditional expectation given $\mathcal{F}_{T-2}$ and using the tower property and (2.4.60) again, with $T-1$ and $T-2$ we have

$$\mathbf{E}\left[U'\left(V_T^{x,\phi^\star}\right)|\mathcal{F}_{T-2}\right] = \mathbf{E}\left[U_{T-1}'\big(x + \sum_{i=1}^{T-2}\langle\phi_i^\star, \Delta S_i\rangle + \left\langle\phi_{T-1}^\star, \Delta S_{T-1}\right\rangle\big)\big|\mathcal{F}_{T-2}\right] = U_{T-2}'(V_{T-2}), \tag{2.4.69}$$

iterating this procedure '$T-t$' times we have

$$\mathbf{E}\left[U_T'\left(V_T^{x,\phi^\star}\right)|\mathcal{F}_t\right] = U_t'\left(V_t^{x,\phi^\star}\right). \tag{2.4.70}$$

In other words, we have the important result

Lemma 2.4.25. *Suppose that the assumptions in Theorem 2.4.16 hold, then there exist a unique optimal strategy $\phi^\star(x)$ and $\left\{U_t'\left(V_t^{x,\phi^\star}\Delta S_t^i\right)\right\}_{t\geqslant 0}$ is a martingale with respect to $\mathbb{P}$ and $\{S_t^i\}_{t\leqslant T}$ is a $\mathbb{Q}-$martingale.*

We explain why the last statement holds. As $\phi^\star(x) = \left\{\tilde{\xi}_t(V_{t-1})\right\}_{1\leqslant t\leqslant T}$ is optimal then applying Proposition 6.6 for any $0\leqslant t\leqslant T$ and $1\leqslant i\leqslant d$, we have $\mathbf{E}\left[U_t'\left(X+\langle\tilde{\xi}_t(X),\Delta S_t\rangle\right)\Delta S_t^i|\mathcal{F}_{t-1}\right] = 0$ for any random variable X, in particular for $X = V_{T-1}^{x,\phi^\star}$ and (2.4.70)

$$\mathbf{E}\left[U_T'\left(V_{T-1}^{x,\phi^\star}+\left\langle\tilde{\xi}_t\left(V_{T-1}^{x,\phi^\star}\right),\Delta S_t\right\rangle\right)\Delta S_t^i|\mathcal{F}_{t-1}\right] = \mathbf{E}\left[U_T'\left(V_T^{x,\phi^\star}\right)\Delta S_T^i|\mathcal{F}_{T-1}\right] = 0. \quad (2.4.71)$$

By well-known properties of conditional expectations and if the probability measure has a density given by (2.4.26) then, by (2.4.71), similarly as explained before $\phi_{T-1}^\star(x) = \tilde{\xi}_{T-1}\left(V_{T-2}^{x,\phi^\star_{T-2}}\right)$ is the optimal 'one-step' strategy at period $T-1$ for the 'initial' (random) endowment $V_{T-2}^{x,\phi^\star}$ then

$$\mathbf{E}\left[U_{T-1}'\left(V_{T-2}^{x,\phi^\star}+\left\langle\tilde{\xi}_{T-1}\left(V_{T-2}^{x,\phi^\star}\right),\Delta S_{T-1}\right\rangle\right)\Delta S_{T-1}^i|\mathcal{F}_{T-2}\right] = 0,$$

again the property of conditional expectation and the martingale property in (2.4.25)

$$\begin{aligned}\mathbf{E}_{\mathbb{Q}}\left[\Delta S_{T-1}^i|\mathcal{F}_{T-2}\right] &= \frac{1}{\mathbf{E}U'\left(V_T^{x,\phi^\star}\right)}\mathbf{E}\left[U_T'\left(V_T^{x,\phi^\star}\right)\Delta S_{T-1}^i|\mathcal{F}_{T-2}\right] = \\ &= \mathbf{E}\left[\mathbf{E}\left[U_T'\left(V_T^{x,\phi^\star}\right)|\mathcal{F}_{T-1}\right]\Delta S_{T-1}^i|\mathcal{F}_{T-2}\right]/\mathbf{E}U'\left(V_T^{x,\phi^\star}\right) =\end{aligned}$$

$$= \frac{1}{\mathbf{E}U'\left(V_T^{x,\phi^\star}\right)}\cdot\mathbf{E}\left[\mathbf{E}\left[U_{T-1}'\left(V_{T-2}^{x,\phi^\star}+\left\langle\tilde{\xi}_{T-1}\left(V_{T-2}^{x,\phi^\star}\right),\Delta S_{T-1}\right\rangle\right)\Delta S_{T-1}^i|\mathcal{F}_{T-2}\right]\right] = 0,$$

for any $1\leqslant t\leqslant T$ and by Proposition 6.4 in [44] (or (2.4.70)) applied to $V_t^{x,\phi^\star}$

$$\mathbf{E}\left[U_{t+1}'\left(V_t^{x,\phi^\star}+\left\langle\tilde{\xi}_{t+1}\left(V_t^{x,\phi^\star}\right),\Delta S_{t+1}\right\rangle\right)\Delta S_{t+1}^i|\mathcal{F}_t\right] = 0,$$

and

$$\begin{aligned}\mathbf{E}_{\mathbb{Q}}\left[\Delta S_{t+1}^i|\mathcal{F}_t\right] &= \mathbf{E}\left[U_T'\left(V_T^{x,\phi^\star}\right)\Delta S_{t+1}^i|\mathcal{F}_t\right]/\mathbf{E}U'\left(V_T^{x,\phi^\star}\right) = \\ &= \mathbf{E}\left[U_{t+1}'\left(V_t^{x,\phi^\star}+\left\langle\tilde{\xi}_{t+1}\left(V_t^{x,\phi^\star}\right),\Delta S_{t+1}\right\rangle\right)\Delta S_{t+1}^i|\mathcal{F}_t\right]/\mathbf{E}U'\left(V_T^{x,\phi^\star}\right) = 0.\end{aligned}$$

In other words, under the measure $\mathbb{Q}$ given by (2.4.26) the process $\{S_t\}_{t\geqslant 0}$ is a martingale with respect to $\mathbb{Q}$. Thus, $\mathbb{Q}$ is a martingale measure (for S).

-

Chapter 3

A DUAL APPROACH TO OPTIMAL INVESTMENT IN DISCRETE-TIME

1 Introduction

Although Expected Utility Theory is one of the basic frameworks to model decision making under uncertainty and an important tool to model economic agents' behaviour their basic principles are debatable, some of its fundamental axioms are inconsistent with empirical evidence on how investors face uncertainty.
In recent years, there has been a burgeoning interest in understanding the behavioural component within decision making processes. Motivated by basic experiments on choices and preferences, Kahnmenan and Tversky developed an alternative set of paradigms that explain investors rationale when facing uncertainty, see [58] and [30]. They conclude that economic agents are subject systematically to biases and error judgment, hence their views on odds are subjective. Investors tend to overweight small probabilities in their favor and underestimate the outcomes that are not favorable. Thus, the axiom leading to a fully rational agent is no longer valid and this, in turn, affects the core of the mathematical model in expected utility, as we cannot assume that investors 'preferred' strategies are based on an actual estimation of how likely some events are, this knowledge is always blurred by investors views.
Kahneman and Tversky, as well as many other economist investigating investors' behaviour have found that even though two agents had the same 'risk aversion profile', their reactions to potential losses and gains may be drastically different, [57]. This led to propose that, apart from their own's biases, each investor has a reference point or benchmark he evaluates outcomes as gains or losses with. Quoting Gilboa I. in [24]: "The distinction between gains and losses based on a reference point is not a violation of an explicit axiom of the classical theory. Rather, it shows that the very language of the classical model, which implicitly assumes that only final outcomes matter, may be too restrictive".

In the present chapter we describe the mathematical setting of Cumulative Prospect Theory (CPT) in a discrete time setting, propose some readily verifiable conditions and prove that if these hold, the problem is well-posed and furthermore this also ensures the existence of optimal strategies. We also relax some assumptions from [11] and hence generalise their models.

The mathematical relevance of the problem is that the functional we aim to optimise is not a traditional one. Therefore, dynamic programing arguments are no longer valid (recall that in order to apply the ideas of backward induction and obtain a optimal strategy the so-called Bellman principle must hold, at least ‘formally’). On the other hand the functionals in consideration may not be concave. Although there exists no natural dual problem for optimisation under behavioural criteria (due to the lack of concavity), we will rely on techniques based on the usual duality between attainable contingent claims and equivalent martingale measures. These results form the base of [42].

The chapter constitutes our first contribution to the theory of optimal investment, we present a series of results that complement and improve results in [11] where the existence of an optimal strategy for an investor with behavioural criteria was proved under certain parameter restrictions (Assumption 3.2.3b below). Here we show the same result under different restrictions on the parameters (Assumption 3.2.3a) but they are neither stronger nor weaker than Assumption 3.2.3b. Assumption 3.2.3a is necessary and sufficient in certain continuous-time models (this is shown in [41] except a borderline case). Furthermore, we manage to reprove the main result of [11] under somewhat weaker assumptions.

The key new ideas are imported from [41] and [44] (some of these ideas were developed in chapter 2) and rely on the construction of an equivalent martingale measure for the price process whose density has nice integrability properties (see Lemma 3.3.1 below) and by Theorem 2.4.16 and Remark 2.4.17 in chapter 2. It is this martingale measure that permits us to prove weakly relative compactness of an optimiser sequence of strategies (Lemma 3.3.12 below).

2 Model description

Fix a real number $T > 0$ acting as time horizon in the sequel and a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0,\ldots,T}\right., \mathbb{P})$ whose initial filtration $\mathcal{F}_0$ is complete.
We consider a financial market evolving in discrete time consisting of d risky assets whose discounted prices are given by an $\mathbb{R}^d$-valued adapted stochastic process, $S = (S_t)_{t=0,\ldots,T}$ where $S_t = \left(S_t^1, \ldots, S_t^d\right)$ for each $t = 0, \ldots, T$.

In addition, we are assuming the financial market to be liquid and frictionless, that is, all costs and constraints associated with transactions are non-existent, investors are allowed to short-sell stocks and to borrow money, and it is always possible to buy or sell an unlimited number of shares of any asset.
As in the last chapter, we denote by Ξ_t^d the set of $d-$dimensional $\mathcal{F}_t-$measurable random variables and $\mathscr{W}$ is the set of $\mathbb{R}$-valued (or $\mathbb{R}^d$-valued) random variables Y such that $\mathbf{E}_{\mathbb{P}} |Y|^p < \infty$ for all $p > 0$.

As explained in chapter 2, trading strategies are characterised by an initial capital z and a d-dimensional process $\{\theta_t : 1 \leqslant t \leqslant T\}$ representing the holdings in the respective assets. We assume θ to be predictable, i.e. $\theta_t \in \Xi_{t-1}^d$ for all t. The class of all such strategies is denoted by Φ.

We define $X_t^z(\theta) := z + \sum_{k=1}^t \langle \theta_k, \Delta S_k \rangle$, the value process of a portfolio with initial investment z and trading strategy θ, where $\Delta S_k := S_k - S_{k-1}$ and $\langle \, , \, \rangle$ denotes scalar product in $\mathbb{R}^d$. For $x \in \mathbb{R}$ the notations x_+, x_- stand for positive and negative parts, respectively.

We shall make the following assumptions throughout this chapter.

Assumption 3.2.1. *For all $t \geqslant 1$, $\Delta S_t \in \mathscr{W}$. Furthermore, for $t = 0, 1, \dots T-1$, there exist $\mathcal{F}_t$-measurable κ_t, $\beta_t > 0$ satisfying $\frac{1}{\kappa_t}$, $\frac{1}{\beta_t} \in \mathscr{W}$ such that*

$$\operatorname*{ess\,inf}_{\xi \in \Xi_t^d} \mathbb{P}\left(\langle \xi, \Delta S_{t+1} \rangle \leqslant -\beta_t |\xi| \,\middle|\, \mathcal{F}_t\right) \geqslant \kappa_t \ a.s. \tag{3.2.1}$$

We may and will assume $\kappa_t, \beta_t \leqslant 1$ in the sequel. As pointed out in [11], (3.2.1) is a strengthened form of the absence of arbitrage condition, Assumption 3.2.1 can be thought as an assumption on the support of the conditional distribution of ΔS_t given $\mathcal{F}_{t-1}$, namely, no assets are 'redundant' and (see chapter 2) $D_t(\omega) = \mathbb{R}^d$ $\mathbb{P} - a.s.$ See [44].
We denote by $\mathcal{M}^e(S)$ the set of equivalent martingale measures for S. Recall that, under the standard no arbitrage hypothesis, $\mathcal{M}^e(S) \neq \varnothing$, see e.g. [27]. Assumption 3.2.1 will allow us to construct a particular $\mathbb{Q} \in \mathcal{M}^e(S)$ with favourable properties, see Lemma 3.3.1 below.

Now we turn to the description of an economic agent. Her attitude towards gains and loses will be described in terms of functions u_+ and u_-. In addition, she will be assumed to distort the 'real world' distributions (probabilities) by means of functions w_+ and w_-. She will further have a 'benchmark' or reference point B which is used when evaluating portfolio payoffs at the terminal time T.

Assumption 3.2.2. *We assume that $u_\pm : \mathbb{R}^+ \to \mathbb{R}^+$ and $w_\pm : [0,1] \to [0,1]$ are measurable functions such that $u_\pm(0) = 0$, $w_\pm(0) = 0$, $w_\pm(1) = 1$, and*

$$u_+(x) \leqslant k_+(x^\alpha + 1), \ \text{for all } x \in \mathbb{R}^+, \tag{3.2.2}$$

$$k_-\left(x^\beta - 1\right) \leqslant u_-(x), \ \text{for all } x \in \mathbb{R}^+, \tag{3.2.3}$$

$$w_+(p) \leqslant g_+ p^\gamma, \ \text{for all } p \in [0,1], \tag{3.2.4}$$

$$w_-(p) \geqslant g_- p^\delta, \ \text{for all } p \in [0,1]. \tag{3.2.5}$$

with $\alpha, \beta, \gamma, \delta > 0$, $k_\pm, g_\pm > 0$ fixed constants.

Assumption 3.2.3. *This concerns the parameters involved in Assumption 3.2.2. For convenience, we shall consider two separate cases.*
Assumption 3.2.3a*. The parameters α, β, γ and δ satisfy*

$$\alpha < \beta \text{ and } \frac{\alpha}{\gamma} < 1 < \frac{\beta}{\delta}. \tag{3.2.6}$$

Assumption 3.2.3a*. The parameters α, β, γ and δ are such that*

$$\frac{\alpha}{\gamma} < \beta. \tag{3.2.7}$$

Assumption 3.2.7 states, intuitively, that the potential growth of the dissatisfaction, u_- must be higher than the 'distorted risk aversion' i.e. the ratio of the growth of the utility u_+ and the distortion on gains, w_+. On the other hand, (3.2.6) in Assumption 3.2.3a states that 'distorted' risk aversion $\frac{\alpha}{\gamma}$ should be lower than the 'distorted' loss aversion $\frac{\beta}{\delta}$, furthermore the dissatisfaction of the outcomes (or the growth condition of such a function) must dominate investor's potential distortions of the likelihood of losses. In a sense, in order to ensure well-posedness, loss aversion should loom both the gains and the distortions on losses. Conversely, distortion on the gains should 'dominate' the utility on gains, that is $\frac{\alpha}{\gamma} < 1$.
Condition (3.2.7) is given in [11], in this chapter we show that this condition is necessary to ensure a well-posed problem and it is valid for a general family of models, more general than those considered in [11]. On the other hand, condition (3.2.6) is proposed in [41]. Under condition (3.2.6) we show that, in discrete-time, few assumptions are required, in particular no market completeness is assumed whatsoever. One can also notice that our conditions (3.2.7) and (3.2.6) are straightforward to check, c.f. Theorem 9.2 in [28].
Now, we discuss our assumptions concerning the reference point.

Assumption 3.2.4. *Similarly to the previous assumption, we consider two cases.*
Assumption 3.2.4a. *The reference point $B \in \Xi_T^1$ belongs to $L^{1+r}(\mathbb{P})$ for some $r > 0$.*

Assumption 3.2.4b. *For the reference point $B \in \Xi_T^1$ there is a trading strategy $\phi \in \Phi$ and initial capital $b \in \mathbb{R}$ satisfying*

$$X_T^b(\phi) = b + \sum_{t=1}^{T} \langle \phi_t, \Delta S_t \rangle \leqslant B. \tag{3.2.8}$$

An economic interpretation can easily be given to Assumption 3.2.4b. The behavioural investor's benchmark are comparable to the value of portfolios. In other words, the benchmark is not low in terms of portfolio values.

Given a real-valued random variable X representing the outcome of an investment, a behavioural agent measures her satisfaction distorting the expected utility of profits as well as the expected 'dissatisfaction' of losses. Consider the nonlinear functionals $V_+(X)$ and $V_-(X)$ defined below. Let

$$V_+(X) := \int_0^\infty w_+(\mathbb{P}(u_+(X_+) > y))\,dy. \tag{3.2.9}$$

Notice that V_+ incorporates the utility of the investor on gains and w_+ produces a non-linear alteration of the given probability distribution. If $w_+(x) = x$ then we return to the expected utility framework since in this case $V_+(X) = \mathbf{E}u_+\left((X-B)_+\right)$. Similarly, let

$$V_-(X) := \int_0^\infty w_-(\mathbb{P}(u_-(X_-) > y))\,dy, \tag{3.2.10}$$

and, finally, the objective or performance functional we aim to optimise is defined by

$$V(X) := V_+(X) - V_-(X), \tag{3.2.11}$$

provided that at least one of the terms is finite.

We define the functionals V_+, V_- below by

$$V_+\left(z,\theta_1,\ldots,\theta_T\right) := V_+\left(X_T^z\left(\theta\right)\right) = \int_0^\infty w_+\left(\mathbb{P}\left(u_+\left(\left(X_T^z\left(\theta\right)-B\right)_+\right) > y\right)\right) dy, \tag{3.2.12}$$

$$V_-\left(z,\theta_1,\ldots,\theta_T\right) := V_-\left(X_T^z\left(\theta\right)\right) = \int_0^\infty w_-\left(\mathbb{P}\left(u_-\left(\left(X_T^z\left(\theta\right)-B\right)_-\right) > y\right)\right) dy. \tag{3.2.13}$$

Definition 3.2.1. We say that a trading strategy $\theta \in \Phi$ is admissible for initial capital z if $V_-\left(X_T^z\left(\theta\right)\right) < \infty$. We denote the set of such trading strategies by $\mathcal{A}\left(z\right)$ and define, for $\theta \in \mathcal{A}(z)$,

$$V\left(z,\theta_1,\ldots,\theta_T\right) := V\left(X_T^z\left(\theta\right)\right) = V_+\left(z,\theta_1,\ldots,\theta_T\right) - V_-\left(z,\theta_1,\ldots,\theta_T\right).$$

Under the tenets of CPT, the functional (3.2.11) is used by investors to assess their satisfaction from a given portfolio at terminal time T.
The optimal portfolio problem for a behavioural investor consists of finding $\theta^\star = \left(\theta_1^\star,\ldots,\theta_T^\star\right)$ such that

$$\sup_{\theta\in\mathcal{A}(z)} V\left(z,\theta_1,\ldots,\theta_T\right) = V\left(z,\theta_1^\star,\ldots,\theta_T^\star\right). \tag{3.2.14}$$

3 Main results

In this section we describe the main results on the problem of optimal investment under behavioural criteria in the setting described above.
First, we justify, using some results explained in chapter 2, the existence of a martingale measure with desirable properties. Secondly, we use some well-known inequalities relating moments of portfolio values and the behavioural functionals described in the last section (see 3.2.11). Such estimates allow to deduce well-posedness of the problem, under Assumptions 3.2.1, 3.2.2, 3.2.3 and 3.2.4 (see below). This enables us to prove some moments estimates on the set of admissible strategies. We deduce some properties of a set of admissible strategies.

As it is well known, most discrete-time market models are incomplete, i.e. $\mathcal{M}^e(S)$ is not a singleton, hence the problem of how to choose a suitable equivalent martingale measure $\mathbb{Q}$ arises. Even though we are exploring the investment under behavioural criteria and there is no concern of finding hedging strategies, we shall exploit the 'richness' of the set $\mathcal{M}^e\left(S\right)$ in order to obtain properties of the laws of admissible strategies $\theta \in \mathcal{A}\left(z\right)$.

Lemma 3.3.1. *Under Assumption 3.2.1 there exists $\mathbb{Q} \in \mathcal{M}^e(S)$ such that for $\rho := d\mathbb{Q}/d\mathbb{P}$ we have both $\rho, 1/\rho \in \mathscr{W}$.*

Proof. We rely on [44] and Theorem 2.4.16, as explained in chapter 2, by means of a suitable utility maximisation problem the existence of a martingale measure with desirable properties can be guaranteed. See (2.4.26) with $U\left(x\right)$ is given by the continuously differentiable, concave function

$$U\left(x\right) = \begin{cases} \frac{1}{2}\left(x-\frac{1}{4}\right)+\frac{1}{8} & \text{if } x \geqslant 0 \\ -\frac{1}{2}\left(x-\frac{1}{2}\right)^2+\frac{1}{8} & \text{if } x<0. \end{cases} \tag{3.3.1}$$

The hypotheses of Theorem 2.4.16 (or [44]) hold by Assumption 3.2.1 and by (3.3.1), hence there is $\mathbb{Q} \in \mathcal{M}^e(S)$ such that

$$\rho = \frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{U'(X_T^0(\phi^*))}{EU'(X_T^0(\phi^*))},$$

for some $\phi^* \in \Phi$. Inspecting the proof of Proposition 7.1 in [44] or Remark 2.4.17 one can easily check that $\phi^*_t \in \mathscr{W}$ for all t. Hence $\rho \in \mathscr{W}$ and ρ is bounded away from 0, a fortiori, $1/\rho \in \mathscr{W}$. □

We fix the probability $\mathbb{Q}$ just constructed for later use. It will be key in establishing moment estimates which underlie our main results. Note also that, under Assumption 3.2.4a, $B \in L^{1+\epsilon}(\mathbb{Q})$ for all $0 < \epsilon < r$ by Hölder's inequality and $\rho \in \mathscr{W}$.

We first address the well-posedness of the optimal portfolio problem for a behavioural investor. We say that the optimal investment problem (3.2.14) is *well-posed* if the supremum in (3.2.14) is finite. If the supremum is infinite then the problem is called *ill-posed*.

We know from section 3 of [11] that $\alpha/\gamma \leqslant \beta/\delta$ and $\alpha < \beta$ are necessary for well-posedness. It is an open problem whether they are sufficient as well. We show below, however, that either (3.2.6) or (3.2.7) are sufficient.

Theorem 3.3.2. *Under Assumptions 3.2.1, 3.2.2, 3.2.3a and 3.2.4a, the optimisation problem (3.2.14) is well-posed. In other words,*

$$\sup_{\theta \in \mathcal{A}(z)} V(z, \theta_1, \ldots, \theta_T) < \infty. \tag{3.3.2}$$

We shall use the auxiliary results given below which were shown in [41] (see Lemmas 3.12, 3.13 and 3.14 there). We include their statements for the sake of completeness.

Lemma 3.3.3. *If a, b and s are positive numbers satisfying $\frac{b}{sa} > 1$ then there exists a constant D such that*

$$\mathbf{E}_{\mathbb{P}}(X^s) \leqslant 1 + D \left(\int_0^\infty \mathbb{P}\left(X^b > y\right)^a dy \right)^{\frac{1}{a}}. \tag{3.3.3}$$

for all non-negative random variables X. ▫

Lemma 3.3.4. *Let $d\mathbb{Q}/d\mathbb{P}, d\mathbb{P}/d\mathbb{Q} \in \mathscr{W}$, $\alpha < \beta$ and $\frac{\alpha}{\gamma} < 1 < \frac{\beta}{\delta}$. Fix $m \in \mathbb{R}$. Then there is some $\eta > 0$ satisfying $\eta < \beta$, $\alpha < \eta$ and $\delta < \eta$, and there exist constants $L_1 = L_1(m)$ and $L_2 = L_2(m)$ such that*

$$\int_0^\infty \mathbb{P}\left((X_+)^\alpha > y\right)^\gamma dy \leqslant L_1 + L_2 \int_0^\infty \mathbb{P}\left((X_-)^\eta > y\right)^\delta dy, \tag{3.3.4}$$

for all random variables X with $\mathbf{E}_{\mathbb{Q}}[X] = m$. ▫

Lemma 3.3.5. *Let a, b and s be strictly positive real numbers such that $s < a < b$ and $s \leqslant 1$. Then there exist $0 < \zeta < 1$ and constants R_1, R_2 such that*

$$\int_0^\infty \mathbb{P}\left(X^a > y\right)^s dy \leqslant R_1 + R_2 \left[\int_0^\infty \mathbb{P}\left(X^b > y\right)^s dy \right]^\zeta, \tag{3.3.5}$$

for all non-negative random variables X. ▫

Remark 3.3.6. Note that in the paper [41] it was assumed that $u_\pm, w_\pm$ are power functions (and not only comparable to power functions as in Assumption 3.2.2 above). Furthermore, $\alpha, \beta, \gamma, \delta \leqslant 1$ were stipulated, in line with the literature. One can check in [41] that the proof of Lemma 3.3.4 above goes through under Assumption 3.2.2 only.

These lemmas allow us to prove Theorem 3.3.2.

Proof of Theorem 3.3.2. We imitate the proof of Theorem 3.15 in [41]. By contradiction, let us suppose that the optimisation problem is ill-posed. Then for a sequence $\phi(n) \in \mathcal{A}(z)$, $n \in \mathbb{N}$ we have $V_+([X_T^z(\phi(n)) - B]_+) \to +\infty$ as $n \to +\infty$. Note that, for any non-negative X,

$$V_+(X) \leqslant \int_0^\infty g_+ \mathbb{P}(X^\alpha > (y/k_+) - 1)^\gamma dy \leqslant \int_0^\infty g_+ k_+ \mathbb{P}(X^\alpha > t)^\gamma dt + g_+ k_+.$$

Thus it follows from Lemma 3.3.4 (with the choice $m := z - \mathbf{E}_{\mathbb{Q}}[B]$) that

$$\lim_{n \to +\infty} \int_0^{+\infty} \mathbb{P}\big([X_T^z(\phi(n)) - B]_-^\eta > y\big)^\delta dy = +\infty$$

for some η satisfying $\eta < \beta$, $\alpha < \eta$ and $\delta < \eta$. Notice that

$$V_-(X) \geqslant \int_0^\infty g_- \mathbb{P}(k_- X^\beta - k_- > y)^\delta dy \geqslant \int_1^\infty g_- k_- \mathbb{P}(X^\beta > t)^\delta dt. \tag{3.3.6}$$

Consequently, we can apply Lemma 3.3.5 to conclude that also

$$\lim_{n \to +\infty} V_-([X_T^z(\phi(n)) - B]_-) = +\infty.$$

Therefore, using Lemmas 3.3.4 and 3.3.5 again (and recalling that $0 < \zeta < 1$),

$$\begin{aligned} & V(X_T^z(\phi(n)) - B) \leqslant g_+ k_+ (L_1 + 1) + g_+ k_+ L_2 \int_0^{+\infty} \mathbb{P}(([X_T^z(\phi(n)) - B]_-)^\eta > y)^\delta dy \\ - \quad & V_-([X_T^z(\phi(n)) - B]_-) \leqslant g_+ k_+ (L_1 + 1 + L_2 R_1) \\ + \quad & g_+ k_+ L_2 R_2 \left[\frac{V_-([X_T^z(\phi(n)) - B]_-)}{g_- k_-} + 1 \right]^\zeta - V_-([X_T^z(\phi(n)) - B]_-) \xrightarrow[n \to +\infty]{} -\infty, \end{aligned}$$

which is absurd. Hence, as claimed, the problem is well-posed. □

We present a result about well-posedness under the alternative conditions Assumptions 3.2.3b and 3.2.4b as well. It is worth pointing out that while the conclusions of Theorems 3.3.2 and 3.3.7 are identical, the methods for proving them are significantly different.

Theorem 3.3.7. *Let $\delta \leqslant 1$. Under Assumptions 3.2.1, 3.2.2, 3.2.3b and 3.2.4b the problem is well-posed, i.e.*

$$\sup_{\theta \in \mathcal{A}(z)} V(z, \theta_1, \dots, \theta_T) < \infty. \tag{3.3.7}$$

Proof. Notice that $\delta \leqslant 1$ and (3.2.5) imply the fourth inequality in Assumption 4.1 of [11]. Hence our result follows from Theorem 4.4 in [11]. Note that in [11] $\alpha, \beta, \gamma \leqslant 1$ were also assumed. As already indicated in Remark 4.2 of [11], the proofs go through without this restriction. □

From now on, the existence of optimal strategies will be our main concern. We will need to assume that the filtration is rich enough in the sense of Assumption 3.3.1 below. This assumption means that investors randomize their strategies or, from a mathematical point of view, that we enlarge the underlying probability space. We will comment on this in section 3.4 as well.

Assumption 3.3.1. *Define $\mathcal{G}_0 = \{\varnothing, \Omega\}$, and $\mathcal{G}_t = \sigma(Z_1, \dots, Z_t)$ for $t = 1, \dots, T$, where the Z_i, $i = 1, \dots, T$ are $\mathbb{R}^N$-valued independent random variables. The random variable S_0 is* *constant, ΔS_t is $\mathcal{G}_t$-adapted and B is $\mathcal{G}_T$-measurable.*

Furthermore, $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{F}_0$, $t \geqslant 0$, where $\mathcal{F}_0 = \sigma(\varepsilon)$ with ε uniformly distributed on $[0,1]$ and independent of $(Z_1, \ldots, Z_T)$.

Remark 3.3.8. The Assumption 3.3.1 clearly implies that $\Delta S_t = f^{(t)}(Z_1, ..., Z_t)$ for some Borel functions $f^{(t)}$, for all t and $B = g_B(Z_1, \ldots, Z_T)$ for some Borel function g_B. We may and will suppose without loss of generality that each of the Z_i is bounded.

In [41] the existence of optimal strategies was shown under Assumption 3.2.3a (and $B \in L^1(\mathbb{Q})$ for some reference probability $\mathbb{Q} \in \mathcal{M}^e(S)$) in a (narrow) class of continuous-time models. In [11] existence was shown under Assumptions 3.2.3b, 3.2.4b and 3.3.1 in discrete-time models assuming also the continuity of $f^{(t)}, g_B$. In the present chapter we shall prove existence of an optimiser in discrete-time models under Assumption 3.3.1 and either Assumption 3.2.3a or Assumption 3.2.3b, and we do not need continuity of $f^{(t)}, g_B$. We first present some preparatory results.

Proposition 3.3.9. *Let Assumptions 3.2.1, 3.2.2, 3.2.3a and 3.2.4a hold and take $\mathbb{Q} \sim \mathbb{P}$ as constructed in Lemma 3.3.1. Further, suppose that a sequence of trading strategies $\{\theta^n\} \subset \mathcal{A}(z)$ satisfies*

$$\sup_n V_-(z, \theta_1^n, \ldots, \theta_T^n) < \infty. \tag{3.3.8}$$

Then there exists $\pi > 1$ such that

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left(X_T^z(\theta^n)\right)_-^\pi < \infty, \tag{3.3.9}$$

and

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left(X_T^z(\theta^n)\right)_+ < \infty. \tag{3.3.10}$$

It follows also that

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left[\sup_{t \leqslant T} \left(X_t^z(\theta^n)\right)_-^\pi\right] < \infty, \tag{3.3.11}$$

$$\sup_{n,t} \mathbf{E}_{\mathbb{Q}} \left[\left|X_t^z(\theta^n)\right|\right] < \infty. \tag{3.3.12}$$

Proof. This is a direct application of Lemma 3.3.5. Indeed, choose $1 < s < \frac{\beta}{\delta}$ and λ such that $1 < \lambda < s < \frac{\beta}{\delta}$. Applying Hölder's inequality,

$$\begin{aligned}\mathbf{E}_{\mathbb{Q}}\left[\left(X_T^z(\theta^n) - B\right)_-^{\frac{s}{\lambda}}\right] = \quad & \mathbf{E}_{\mathbb{Q}}\left[\rho^{1/\lambda}\tfrac{1}{\rho^{1/\lambda}}\left(X_T^z(\theta^n) - B\right)_-^{\frac{s}{\lambda}}\right] \quad \leqslant \\ & \leqslant C\mathbf{E}_{\mathbb{P}}\left[\left(X_T^z(\theta^n) - B\right)_-^s\right]^{1/\lambda},\end{aligned}$$

where $C = \mathbf{E}_{\mathbb{Q}}\left[\rho^{q/\lambda}\right]^{1/q} < \infty$ and q is the conjugate number of λ. Lemma 3.3.3 yields that, for all n,

$$C\mathbf{E}_{\mathbb{P}}\left[\left(X_T^z(\theta^n) - B\right)_-^s\right]^{1/\lambda} \leqslant C\left(1 + D\left(\int_0^\infty \mathbb{P}\left(\left(X_T^z(\theta^n) - B\right)_-^\beta > y\right)^\delta dy\right)^{1/\delta}\right)^{1/\lambda}. \tag{3.3.13}$$

for some $0 < D < \infty$. Hence (3.3.8) and (3.3.6) imply (3.3.9), setting $\pi := \min\{\frac{s}{\lambda}, 1 + \frac{r}{2}\}$ (note that, as we have pointed out after Lemma 3.3.1, $\mathbf{E}_{\mathbb{Q}}|B|^{1+(r/2)} < \infty$). Moreover, Hölder's inequality gives

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left(X_T^z\left(\theta^n\right)\right)_- < \infty.$$

It follows from Theorem 2.3.8, that $\left\{X_t^z\left(\theta^n\right)\right\}_{t \leqslant T}$ is a martingale under $\mathbb{Q}$, thus

$$\mathbf{E}_{\mathbb{Q}} \left|X_t^z\left(\theta^n\right)\right| \leqslant \mathbf{E}_{\mathbb{Q}} \left|X_T^z\left(\theta^n\right)\right|,$$

for all n, t. From $\mathbf{E}_{\mathbb{Q}}\left[X_t^z\left(\theta^n\right)\right] = z$ and (3.3.9) we have

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left(X_T^z\left(\theta^n\right)\right)_+ \leqslant |z| + \sup_n \mathbf{E}_{\mathbb{Q}} \left(X_T^z\left(\theta^n\right)\right)_- < \infty.$$

Hence $\sup_n \mathbf{E}_{\mathbb{Q}} \left|X_T\left(\theta^n\right)\right| < \infty$ and this implies (3.3.10) as well as (3.3.12). In order to prove (3.3.11), Doob's inequality is applied, noting that $f(x) = x_-$ is convex and hence the process $\left\{\left(X_t^z\left(\theta^n\right)\right)_-\right\}_{t \leqslant T}$ is a positive submartingale. ☐

Notice that we could show Proposition 3.3.9 only in a discrete time and finite horizon setting since it relies on Theorem 2 of [27], i.e. Theorem 2.3.8, which fails in other (e.g. continuous-time) settings.

Remark 3.3.10. In [41] admissible strategies θ were required to satisfy *both* $V_-(X_T^z(\theta)) < \infty$ and the martingale property for $X_t^z(\theta)$ (under some fixed $\mathbb{Q} \in \mathcal{M}^e(S)$). The proof above shows that, in the present discrete-time setting, $V_-(X_T^z(\theta)) < \infty$ *implies* the martingale property for $X_t^z(\theta)$ under $\mathbb{Q}$. So the domain of optimisation in the present work is the same as the one in [41].

Remark 3.3.11. Let $\theta = (\theta_1, \theta_2, \ldots, \theta_T) \in \mathcal{A}(z)$ be as in Proposition 3.3.9. Clearly,

$$\left(\langle \theta_T, \Delta S_T \rangle\right)_+ \leqslant \left(X_T^z(\theta)\right)_+ + \left(X_{T-1}^z(\theta)\right)_-. \tag{3.3.14}$$

Thus

$$\mathbf{E}_{\mathbb{Q}} \left(\langle \theta_T, \Delta S_T \rangle\right)_+ \leqslant \mathbf{E}_{\mathbb{Q}} \left(X_T^z(\theta)\right)_+ + \mathbf{E}_{\mathbb{Q}} \left(X_{T-1}^z(\theta)\right)_-$$

implies

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left[\left(\langle \theta_T^n, \Delta S_T \rangle\right)_+\right] < \infty. \tag{3.3.15}$$

We will now proceed to proving that trading strategies satisfying (3.3.8) have some uniformly bounded moments.

Lemma 3.3.12. *Let $\{\theta^n\} \subset \mathcal{A}(z)$ be a sequence of trading strategies. Let Assumptions 3.2.1, 3.2.2, 3.2.3a and 3.2.4a hold and assume that* (3.3.8) *holds. Then*

$$\sup_n \mathbf{E}_{\mathbb{Q}} \left|\theta_t^n\right|^{1/2} < \infty \text{ for } t = 1, 2, \ldots, T. \tag{3.3.16}$$

Proof. A uniform bound for $\mathbf{E}_{\mathbb{Q}}\left[\left(\langle \theta_T^n, \Delta S_T \rangle\right)_+\right]$ can obtained as in Remark 3.3.11. Using the same idea for $t \leqslant T$,

$$\begin{aligned} \mathbf{E}_{\mathbb{Q}} \langle \theta_t^n, \Delta S_t \rangle_+ &\leqslant \mathbf{E}_{\mathbb{Q}} \left(X_t^z\left(\theta^n\right)\right)_+ + \mathbf{E}_{\mathbb{Q}} \left(X_{t-1}^z\left(\theta^n\right)\right)_- \leqslant \\ &\leqslant \mathbf{E}_{\mathbb{Q}} \left|X_t^z\left(\theta^n\right)\right| + \mathbf{E}_{\mathbb{Q}} \left|X_{t-1}^z\left(\theta^n\right)\right|, \end{aligned} \tag{3.3.17}$$

and the right-hand side is bounded uniformly in n by Proposition 3.3.9.

Denote by $\rho_t := \mathbf{E}_{\mathbb{P}}[\rho|\mathcal{F}_t]$. From Assumption 3.2.1,

$$\begin{aligned}\mathbf{E}_{\mathbb{Q}}\left\langle\theta_T^n,\Delta S_T\right\rangle_+ &\geqslant \mathbf{E}_{\mathbb{Q}}\left[|\theta_T^n|\left\langle\frac{\theta_T^n}{|\theta_T^n|},\Delta S_T\right\rangle\mathbf{1}_{\left\{\left\langle\frac{\theta_T^n}{|\theta_T^n|},\Delta S_T\right\rangle\geqslant\beta_{T-1}\right\}}\right]\geqslant & (3.3.18)\\ &\geqslant \mathbf{E}_{\mathbb{Q}}\left[|\theta_T^n|\,\beta_{T-1}\mathbb{Q}\left(\left\langle\frac{\theta_T^n}{|\theta_T^n|},\Delta S_T\right\rangle\geqslant\beta_{T-1}\middle|\mathcal{F}_{T-1}\right)\right]. & (3.3.19)\end{aligned}$$

Using the property

$$\mathbf{E}_{\mathbb{Q}}\left[\eta|\,\mathcal{G}\right]=\mathbf{E}_{\mathbb{P}}\left[\eta\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\,\mathcal{G}\right]/\mathbf{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\,\mathcal{G}\right],$$

of conditional expectations which holds for any sigma-algebra $\mathcal{G}$ and for any positive random variable η, we get

$$\mathbf{E}_{\mathbb{Q}}\left\langle\theta_T^n,\Delta S_T\right\rangle_+\geqslant\mathbf{E}_{\mathbb{Q}}\left[\frac{|\theta_T^n|}{\rho_{T-1}}\beta_{T-1}\mathbf{E}_{\mathbb{P}}\left[\mathbf{1}_{\left\{\left\langle\frac{\theta_T^n}{|\theta_T^n|},\Delta S_T\right\rangle\geqslant\beta_{T-1}\right\}}\rho_T\middle|\,\mathcal{F}_{T-1}\right]\right]. \quad (3.3.20)$$

Denote $A_T=\left\{\left\langle\frac{\theta_T^n}{|\theta_T^n|},\Delta S_T\right\rangle\geqslant\beta_{T-1}\right\}$ and apply the (conditional) Cauchy-Schwarz inequality to the right-hand side:

$$\mathbf{E}_{\mathbb{P}}\left[\mathbf{1}_{A_T}\rho_T|\,\mathcal{F}_{T-1}\right]\geqslant\mathbb{P}^2\left(A_T\,|\mathcal{F}_{T-1}\right)\Big/\mathbf{E}_{\mathbb{P}}\left[\frac{1}{\rho_T}\middle|\,\mathcal{F}_{T-1}\right].$$

From Assumption 3.2.1 and Cauchy-Schwarz's inequality, the right-hand side of (3.3.20) can be minorised by

$$\begin{aligned}&\mathbf{E}_{\mathbb{Q}}\left[\frac{|\theta_T^n|}{\rho_{T-1}}\beta_{T-1}\kappa_{T-1}^2/\mathbf{E}_{\mathbb{P}}\left[\rho_T^{-1}\middle|\,\mathcal{F}_{T-1}\right]\right]\geqslant\\ &\qquad\mathbf{E}_{\mathbb{Q}}\left[|\theta_T^n|^{1/2}\right]^2/\mathbf{E}_{\mathbb{Q}}\left[\rho_{T-1}\beta_{T-1}^{-1}\kappa_{T-1}^{-2}\mathbf{E}_{\mathbb{P}}\left[\rho_T^{-1}\middle|\,\mathcal{F}_{T-1}\right]\right], \quad (3.3.21)\end{aligned}$$

thus

$$\mathbf{E}_{\mathbb{Q}}\left\langle\theta_T,\Delta S_T\right\rangle_+\cdot\mathbf{E}_{\mathbb{Q}}\left[\rho_{T-1}\beta_{T-1}^{-1}\kappa_{T-1}^{-2}\mathbf{E}_{\mathbb{P}}\left[\rho_T^{-1}\middle|\,\mathcal{F}_{T-1}\right]\right]\geqslant\mathbf{E}_{\mathbb{Q}}\left[|\theta_T|^{1/2}\right]^2.$$

The same procedure applies to θ_t, $t=1,\dots,T$ by (3.3.17). Thus, for all $t\leqslant T$,

$$\sup_n\mathbf{E}_{\mathbb{Q}}\left[|\theta_t^n|^{1/2}\right]\leqslant\sup_n\left[\mathbf{E}_{\mathbb{Q}}\left\langle\theta_t^n,\Delta S_t\right\rangle_+\mathbf{E}_{\mathbb{Q}}\left[\rho_{t-1}\beta_{t-1}^{-1}\kappa_{t-1}^{-2}\mathbf{E}_{\mathbb{P}}\left[\rho_t^{-1}\middle|\,\mathcal{F}_{t-1}\right]\right]\right]^{1/2}<\infty,$$

by Assumption 3.2.1 and (3.3.17). □

Remark 3.3.13. Applying Hölder's inequality, the estimates above can be carried out with no significant alteration for any $0<\xi<1$, i.e.

$$\sup_n\mathbf{E}_{\mathbb{Q}}\left|\theta_t^n\right|^{\xi}<\infty \text{ for } t=1,2,\dots,T. \quad (3.3.22)$$

can be shown. For simplicity we did this only for $\xi=\frac{1}{2}$.

From Lemma 3.3.12 the next one follows trivially.

Lemma 3.3.14. *Under Assumptions 3.2.1, 3.2.2, 3.2.3a and 3.2.4a, let* $\{\theta^n\}_{n\geqslant1}\subset\mathcal{A}(z)$ *a* *sequence of admissible trading strategies such that* $\sup_n V_-\left(z, \theta_1^n, \dots \theta_T^n\right) < \infty$. *Then* $\{\theta^n\}_{n\geqslant 1}$ *is a tight sequence of* $\mathbb{R}^{dT}$*−valued random variables on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$.

Proposition 3.3.15. *Let* $\{\theta^n\}_{n\geqslant 1}$ *be a sequence of trading strategies whose set of laws is tight. Let* μ_n *be the law of* $X_T^z(\theta^n) - B$ *for all* n. *Under Assumption 3.3.1, there exists a law* $\mu^\star$ *and a trading strategy* $\theta^\star$ *such that* $\mu^\star = Law\left(X_T^z(\theta^\star) - B\right)$ *and* $\mu^\star$ *is an accumulation point of the sequence* $\{\mu_n\}_{n\geqslant 1}$ *in the weak (narrow) topology.*

Proof. Lemma 2.2.7 provides independent random variables $\varepsilon', \tilde{\varepsilon}$, uniform on $[0,1]$, which are both functions of ε. Now following the proof of Theorem 6.8 of [11] (with $\phi_1 = \ldots = \phi_T = 0$). By Lemma 2.2.7 with $l = 2$ applied to ε it is possible to obtain ε' and $\tilde{\varepsilon}$ independent and uniformly distributed on $[0,1]$ and $\mathcal{F}_0$-measurable. By the tightness assumption and by Lemma 2.2.5 the sequence of laws of the random variables

$$\left\{(\varepsilon', \theta_1^n, \dots, \theta_T^n, Z_1, \dots, Z_T)\right\}_{n\geqslant 1},$$

is tight and admits an accumulation point $\mu^\star$ in the weak topology. Denote by M a random variable with law μ. Notice that M takes values in $\mathbb{R}^{T(N+d)+1}$, let μ_k be the marginal of $\mu^\star$ with respect to its first $1+kd$ and last NT coordinates. Lemma 2.2.7 allows to have $\sigma(\tilde{\varepsilon})$-measurable random variables $\varepsilon_1, \dots, \varepsilon_T$ that are independent and uniformly distributed on $[0,1]$. Applying Lemma 2.2.8 with $Y = \varepsilon'$ and $U = \varepsilon_1$ we obtain a function G such that $(\varepsilon', G(\varepsilon', \varepsilon_1))$ has the same law as μ_1. Notice that if $Q = \left(M^1, \dots, M^{d+1}\right)$, $Q' = (\varepsilon', G(\varepsilon', \varepsilon_1))$ and $U = \left(M^{1+dt+1}, \dots, M^{1+d(T+N)}\right)$, $U' = (Z_1, \dots, Z_T)$ then (Q', U') and (Q, U) have the same law as Q and U are independent and so they are Q' and U'. Define $\theta_1^\star := G(\varepsilon', \varepsilon_1)$ and notice that it is $\mathcal{F}_0$-measurable. Proceeding inductively, suppose $(\varepsilon', \theta_1^\star, \dots, \theta_k^\star, Z_1, \dots, Z_T)$ has a law equal to μ_k and $\theta_j^\star$ is a function of ε', $Z_1, \dots, Z_{j-1}$, $\varepsilon_1, \dots, \varepsilon_j$ for any $j = 1, 2\dots, k$. Applying Lemma 2.2.8 with $n_1 = d$, $n_2 = kd+1$ to $U = \varepsilon_{k+1}$ and

$$Y_k = \left(\varepsilon', \theta_1^\star, \dots, \theta_k^\star, Z_1, \dots, Z_k\right),$$

ensures that there is a measurable function G such that $(Y_k, G(Y_k, \varepsilon_{k+1}))$ has the same law as

$$\left(M^1, \dots, M^{1+dk}, M^{1+Td+1}, \dots, M^{1+Td+kN}, M^{1+kd+1}, \dots, M^{1+(k+1)d}\right),$$

then we can take $\theta_{k+1} = G(Y_k, \varepsilon_{k+1})$. Iterating this argument, we obtain a random variable $Y := (\varepsilon', \theta_1^\star, \dots, \theta_T^\star, Z_1, \dots, Z_T)$ having the law $\mu^\star$. Hence, we obtain an $\mathcal{F}_t$-predictable process $\theta_t^\star$ such that, by Prokhorov's theorem, Theorem 2.2.3, the law of

$$Y_k := (\varepsilon', \theta_1^{n_k}, \dots, \theta_T^{n_k}, Z_1, \dots, Z_T),$$

tends to that of

$$Y := (\varepsilon', \theta_1^\star, \dots, \theta_T^\star, Z_1, \dots, Z_T),$$

for a subsequence n_k, as $k \to \infty$.
Skorokhod's theorem provides random variables,

$$\bar{Y}_k = (\bar{\varepsilon}'(n_k), \bar{\theta}_1^{n_k}, \dots, \bar{\theta}_T^{n_k}, \bar{Z}_1^{n_k}, \dots, \bar{Z}_T^{n_k})$$

and

$$\bar{Y} = (\bar{\varepsilon}', \bar{\theta}_1^\star, \dots, \bar{\theta}_T^\star, \bar{Z}_1, \dots, \bar{Z}_T),$$

on some probability space such that $Law(Y_k) = Law(\bar{Y}_k)$ for all k, $Law(Y) = Law(\bar{Y})$ and $\bar{Y}_k$ tends to $\bar{Y}$ a.s.

By assumption, we also have that $\Delta S_i = f^{(i)}\left(Z_1, \dots, Z_i\right)$ and $B = g_B(Z_1, \dots, Z_T)$. Denoting $\Delta \bar{S}_i = f^{(i)}\left(\bar{Z}_1, \dots, \bar{Z}_i\right)$ and $\bar{B} := g_B(\bar{Z}_1, \dots, \bar{Z}_T)$ we have that

$$Law\left((Z_i)_{i\leqslant T}, (\theta_i^\star)_{i\leqslant T}, (\Delta S_i)_{i\leqslant T}, B\right) = Law\left(\left(\bar{Z}_i\right)_{i\leqslant T}, \left(\bar{\theta}_i\right)_{i\leqslant T}, \left(\Delta \bar{S}_i\right)_{i\leqslant T}, \bar{B}\right).$$

Therefore

$$Law\left(\sum_{i=1}^T \left\langle \theta_i^\star, \Delta S_i \right\rangle - B\right) = Law\left(\sum_{i=1}^T \left\langle \bar{\theta}_i, \Delta \bar{S}_i \right\rangle - \bar{B}\right). \tag{3.3.23}$$

Denote $\bar{B}^k := g_B(\bar{Z}_1^{n_k}, \dots, \bar{Z}_T^{n_k})$ and $\Delta \bar{S}_i^k := f^{(i)}(\bar{Z}_1^{n_k}, \dots, \bar{Z}_i^{n_k})$. By Theorem 2.2.6 or Théoréme 1 in [3], $\Delta \bar{S}_i^k \to \Delta \bar{S}_i$ for all i in probability and also $\bar{B}^k \to \bar{B}$ in probability, $k \to \infty$.

It follows that

$$\sum_{i=1}^T \left\langle \bar{\theta}_i^k, \Delta \bar{S}_i^k \right\rangle - \bar{B}^k \to \sum_{i=1}^T \left\langle \bar{\theta}_i, \Delta \bar{S}_i \right\rangle - \bar{B}, \tag{3.3.24}$$

in probability, hence in law. In other words, we have

$$\sum_{i=1}^T \left\langle \theta_i^k, \Delta S_i \right\rangle - B \to \sum_{i=1}^T \left\langle \theta_i^\star, \Delta S_i \right\rangle - B, \tag{3.3.25}$$

in law. This finishes the proof. □

Our first main result on the existence of optimal strategies now follows easily from Proposition 3.3.15 above.

Main Theorem 3.3.16. *Let Assumptions 3.3.1, 3.2.1, 3.2.2, 3.2.3a and 3.2.4a be in force and let $u_\pm$, $w_\pm$ be continuous. Then the supremum in (3.2.14) is attained by an optimal strategy $\theta^\star$.*

Proof. Let us take a maximising sequence of admissible strategies $\left\{\theta^j\right\}_{j\geqslant 1}$, i.e.

$$V\left(z, \theta_1^j, \dots, \theta_T^j\right) \to \sup_{\theta\in\mathcal{A}(z)} V\left(z, \theta_1, \dots, \theta_T\right),\ j \to \infty.$$

The proof of Theorem 3.3.2 shows that we necessarily have $\sup_j V_-\left(z, \theta_1^j, \dots, \theta_T^j\right) < \infty$, showing (3.3.8).

Due to Lemma 3.3.14, we can conclude that the sequence $\left\{\theta^j\right\}_{j\geqslant 1}$ is tight. Proposition 3.3.15 then shows that there is a strategy $\theta^\star \in \mathcal{A}(z)$, $\theta^\star = (\theta_1^\star, \theta_2^\star, \dots, \theta_T^\star) \in \mathbb{R}^{dT}$ such that $X_T^z\left(\theta^j\right) - B \to X_T^z\left(\theta^\star\right) - B$ in law (along a subsequence which we assume to be the original sequence). Now our aim is to prove

$$\limsup_j V\left(z, \theta_1^j, \theta_2^j, \dots, \theta_T^j\right) \leqslant V\left(z, \theta_1^\star, \theta_2^\star, \dots, \theta_T^\star\right). \tag{3.3.26}$$

By the continuous mapping theorem, Theorem 2.1.10, we have

$$\left(X_T^z\left(\theta^j\right) - B\right)_\pm \to \left(X_T^z\left(\theta^\star\right) - B\right)_\pm,$$

and

$$u_{\pm}\left(\left(X_T^z\left(\theta^j\right)-B\right)_{\pm}\right) \to u_{\pm}\left(\left(X_T^z\left(\theta^\star\right)-B\right)_{\pm}\right), \tag{3.3.27}$$

in law. Let D be the set of discontinuity points of the limiting distributions in (3.3.27). Then

$$\mathbb{P}\left(u_{\pm}\left(\left(X_T^z\left(\theta^j\right)-B\right)_{\pm}\right) \geqslant y\right) \to \mathbb{P}\left(u_{\pm}\left(\left(X_T^z\left(\theta^\star\right)-B\right)_{\pm}\right) \geqslant y\right) \quad \text{for all } y \in \mathbb{R}^+ \backslash D,$$

in particular, for Lebesgue-a.e. y. By Assumption 3.2.2,

$$w_+\left(\mathbb{P}\left(u_+\left(\left(X_T^z\left(\theta^j\right)-B\right)_+\right) \geqslant y\right)\right) \leqslant g_+\left[\mathbb{P}\left(u_+\left(\left(X_T^z\left(\theta^j\right)-B\right)_+\right) \geqslant y\right)\right]^{\gamma}. \tag{3.3.28}$$

Take $1/\gamma < \lambda < 1/\alpha$. Applying Markov's inequality and Assumption 3.2.2 again,

$$g_+\left[\mathbb{P}\left(u_+^{\lambda}\left(\left(X_T^z\left(\theta^j\right)-B\right)_+\right) \geqslant y^{\lambda}\right)\right]^{\gamma} \leqslant \frac{c'}{y^{\lambda\gamma}}\left\{\mathbf{E}_{\mathbb{P}}\left(1+\left(X_T^z\left(\theta^j\right)-B\right)_+^{\alpha}\right)^{\lambda}\right\}^{\gamma}, \tag{3.3.29}$$

for some $c' > 0$, hence

$$c'\left[\mathbb{P}\left(u_+^{\lambda\gamma}\left(\left(X_T^z\left(\theta^j\right)-B\right)_+\right) \geqslant y^{\lambda}\right)\right]^{\gamma} \leqslant \frac{c''}{y^{\lambda\gamma}}\mathbf{E}_{\mathbb{P}}^{\gamma}\left(1+\left(X_T^z\left(\theta^j\right)-B\right)_+^{\alpha\lambda}\right),$$

with some $c'' > 0$. Furthermore,

$$\mathbf{E}_{\mathbb{P}}^{\gamma}\left(1+\left(X_T^z\left(\theta^j\right)-B\right)_+^{\alpha\lambda}\right) \leqslant c'''\left(1+\left[\mathbf{E}_{\mathbb{P}}\left(X_T^z\left(\theta^j\right)\right)_+^{\alpha\lambda}\right]^{\gamma}+\left[\mathbf{E}_{\mathbb{P}}B_-^{\alpha\lambda}\right]^{\gamma}\right). \tag{3.3.30}$$

The last term is finite by Assumption 3.2.4a. Hölder's inequality applied with $p = \frac{1}{\alpha\lambda}$ gives

$$\mathbf{E}_{\mathbb{P}}\left(X_T^z\left(\theta^j\right)\right)_+^{\alpha\lambda} = \mathbf{E}_{\mathbb{Q}}\left[\frac{1}{\rho}\cdot\left(X_T^z\left(\theta^j\right)\right)_+^{\alpha\lambda}\right] \leqslant C_1\left[\mathbf{E}_{\mathbb{Q}}\left(X_T^z\left(\theta^j\right)\right)_+\right]^{\alpha\lambda},$$

with $C_1 = \mathbf{E}_{\mathbb{Q}}\left[\rho^{1/(\alpha\lambda-1)}\right]^{1-\alpha\lambda}$. Thus

$$w_+\left(\mathbb{P}\left(u_+\left(\left(X_T^z\left(\theta^j\right)-B\right)_+\right) \geqslant y\right)\right) \leqslant \frac{1}{y^{\lambda\gamma}}\left\{D_1+D_2\left[\mathbf{E}_{\mathbb{Q}}\left(X_T^{x_0}\left(\theta^j\right)\right)_+\right]^{\alpha\lambda\gamma}\right\},$$

with suitable constants D_1, D_2.
The condition $\sup_j V_-\left(z,\theta_1^j,\ldots,\theta_T^j\right) < \infty$ implies that $\sup_j \mathbf{E}_{\mathbb{Q}}\left(X_T^z\left(\theta^j\right)\right)_+ < \infty$ (see Proposition 3.3.9). This in turn gives that the sequence of positive functions

$$w_+\left(\mathbb{P}\left(u_+\left(\left(X_T^z\left(\theta^j\right)-B\right)_+\right) \geqslant y\right)\right),$$

can be dominated by $\frac{K}{y^{\lambda\gamma}}$ for some $K > 0$.
This estimate allows to apply Fatou's lemma since

$$w_+\left(\mathbb{P}\left(u_+\left(\left(X^z\left(\theta^j\right)-B\right)_+\right) \geqslant y\right)\right) \leqslant \mathbf{I}_{[0,1]}(y) + \frac{K}{y^{\lambda\gamma}}\cdot\mathbf{I}_{(1,\infty)}(y), \tag{3.3.31}$$

which yields

$$\limsup_j V_+\left(z,\theta_1^j,\ldots,\theta_T^j\right) \leqslant \int_0^{\infty}\limsup_j w_+\left(\mathbb{P}\left(u_+\left(\left(X^z\left(\theta^j\right)-B\right)_+\right) \geqslant y\right)\right)dy,$$

and the latter equals $V_+ (z, \theta_1^\star, \ldots, \theta_T^\star)$.

On the other hand, by Fatou's lemma applied to $V_- \left(z, \theta_1^j, \ldots, \theta_T^j\right)$ we have

$$V_- (z, \theta_1^\star, \ldots, \theta_T^\star) \leqslant \liminf_j V_- \left(z, \theta_1^j, \ldots, \theta_T^j\right) \leqslant \sup_j V_- \left(x_0, \theta_1^j, \ldots, \theta_T^j\right) < \infty,$$

so $\theta^\star \in \mathcal{A}(z)$ and

$$\limsup_j V^+ \left(z, \theta_1^j, \ldots, \theta_T^j\right) - \liminf_j V^- \left(z, \theta_1^j, \ldots, \theta_T^j\right) \leqslant$$
$$V^+ (z, \theta_1^\star, \ldots, \theta_T^\star) - V^- (z, \theta_1^\star, \ldots, \theta_T^\star),$$

thus

$$\limsup_j \left\{V^+ \left(z, \theta_1^j, \ldots, \theta_T^j\right) - V^- \left(z, \theta_1^j, \ldots, \theta_T^j\right)\right\} \leqslant$$
$$V^+ (z, \theta_1^\star, \ldots, \theta_T^\star) - V^- (z, \theta_1^\star, \ldots, \theta_T^\star),$$

which yields (3.3.26). Hence $\theta^\star$ is optimal and the supremum is attainable. □

Now we turn to the case of Assumptions 3.2.3b and 3.2.4b.

Main Theorem 3.3.17. *Let Assumptions 3.3.1, 3.2.1, 3.2.2, 3.2.3b and 3.2.4b be in force, assume $\delta \leqslant 1$ and let $u_\pm$, $w_\pm$ be continuous. Then, the supremum is attained in* (3.2.14) *by some $\theta^\star \in \mathcal{A}(z)$.*

Proof. We can follow the proof of Theorem 6.8 in [11] verbatim up to the point of constructing

$$Y := (\varepsilon', \theta_1^\star, \ldots, \theta_T^\star, Z_1, \ldots, Z_T).$$

Then the argument of Proposition 3.3.15 shows that

$$\sum_{i=1}^{T} \left\langle \theta_i^k, \Delta S_i \right\rangle - B \to \sum_{i=1}^{T} \left\langle \theta_i^\star, \Delta S_i \right\rangle - B, \tag{3.3.32}$$

in law. From this point on the Fatou-lemma argument as in Theorem 3.3.16 or Theorem 6.8 in [11] applies and optimality of $\theta^\star$ can be established. □

4 A sufficient condition

Assumption 3.3.1 may look restrictive at first sight. Hence we provide a simple sufficient condition for its validity.

Proposition 3.4.1. *Let $\mathcal{G}_t := \sigma(\tilde{Z}_1, \ldots, \tilde{Z}_t)$ for $t = 1, \ldots, T$ and $\mathcal{G}_0 = \{\varnothing, \Omega\}$ where the $\tilde{Z}_i$, $i = 1, \ldots, T$ are N-dimensional random variables with a Lebesgue-a.e. positive joint density on $\mathbb{R}^{TN}$. Then there are independent $\mathbb{R}^N$-valued random variables Z_i, $i = 1, \ldots, T$ such that $\mathcal{G}_t = \sigma(Z_1, \ldots, Z_t)$ for $t = 1, \ldots, T$.*

We first recall a well-known lemma about the simulation of random variables (see Lemma 9.6 of [11]).

Lemma 3.4.2. *Let X be a real-valued random variable with atomless law. Let $F(x) := P(X \leqslant x)$ denote its cumulative distribution function. Then $F(X)$ has uniform law on $[0,1]$.* ▫

The following results are parallel to Lemma 9.8 and Corollary 9.9 of [11]. In the sequel, when we write "measurable bijection" we mean that both the function and its inverse are measurable.

Lemma 3.4.3. *Let (Y, W) be an $\mathbb{R} \times \mathbb{R}^k$-valued random variable with Lebesgue almost everywhere positive density $f(x^1, \ldots, x^{k+1})$. Then there is a measurable bijection H from $\mathbb{R}^{k+1}$ into $[0,1] \times \mathbb{R}^k$ such that $H^i(x^1, \ldots, x^{k+1}) = x^i$ for $i = 2, \ldots, k+1$ and $Z := H^1(Y, W)$ is uniform on $[0,1]$, independent of W.*

Proof. The conditional distribution function of Y knowing $W = (x^2, \ldots, x^{k+1})$,

$$F(x^1, \ldots, x^{k+1}) := \frac{\int_{-\infty}^{x^1} f(z, x^2, \ldots, x^{k+1}) dz}{\int_{-\infty}^{\infty} f(z, x^2, \ldots, x^{k+1}) dz},$$

is clearly measurable (in all its variables). By a.e. positivity of f, F is also strictly increasing in x^1 hence the function

$$H : (x^1, \ldots, x^{k+1}) \to (F(x^1, \ldots, x^{k+1}), x^2, \ldots, x^{k+1}),$$

is a measurable bijection. By Lemma 3.4.2 the conditional law $P(H^1(Y, W) \in \cdot \,| W = (x^2, \ldots, x^{k+1}))$ is uniform on $[0,1]$ for Lebesgue-almost all $(x^2, \ldots, x^{k+1})$, which shows that $H^1(Y, W)$ is independent of W with uniform law on $[0,1]$. □

Corollary 3.4.4. *Let $(\tilde{W}_1, \ldots, \tilde{W}_k)$ be an $\mathbb{R}^k$-valued random variable with a.e. positive density (w.r.t. the k-dimensional Lebesgue measure). Then there are independent random variables $W_1, \ldots, W_k$ and measurable bijections $g_l(k) : \mathbb{R}^l \to \mathbb{R}^l$, $1 \leqslant l \leqslant k$ such that $(\tilde{W}_1, \ldots, \tilde{W}_l) = g_l(k)(W_1, \ldots, W_l)$.*

Proof. The case $k = 1$ is vacuous. Assume that the statement is true for $k \geqslant 1$, let us prove it for $k+1$. We may set $g_l(k+1) := g_l(k)$, $1 \leqslant l \leqslant k$, it remains to construct $g_{k+1}(k+1)$ and W_{k+1}.

We wish to apply Lemma 3.4.3 in this induction step. It provides a measurable bijection $s : \mathbb{R}^{k+1} \to \mathbb{R}^{k+1}$ such that $s^m(x^1, \ldots, x^{k+1}) = x^m$, $1 \leqslant m \leqslant k$ and $W_{k+1} := s^{k+1}(\tilde{W}_1, \ldots, \tilde{W}_{k+1})$ is independent of $(\tilde{W}_1, \ldots, \tilde{W}_k)$ and hence of

$$(W_1, \ldots, W_k) = g_k(k)^{-1}(\tilde{W}_1, \ldots, \tilde{W}_k).$$

Define $a : \mathbb{R}^{k+1} \to \mathbb{R}^{k+1}$ by

$$\begin{aligned} a(x^1, \ldots, x^{k+1}) &:= (g_k(k)^{-1}(x^1, \ldots, x^k), s^{k+1}(x^1, \ldots, x^{k+1})) \\ &= s(g_k(k)^{-1}(x^1, \ldots, x^k), x^{k+1}), \end{aligned}$$

a is clearly a measurable bijection. Notice that $a(\tilde{W}_1, \ldots, \tilde{W}_{k+1}) = (W_1, \ldots, W_{k+1})$. Set $g_{k+1}(k+1) := a^{-1}$. This finishes the proof of the induction step and hence concludes the proof. □

Proof of Proposition 3.4.1. Apply Corollary 3.4.4 with the choice $k := TN$ and

$$\tilde{W}_{(t-1)N+l} := \tilde{Z}_t^l, \quad l = 1, \ldots, N.$$

and $t = 1, \ldots, T$. By the construction in Corollary 3.4.4, taking $Z_t^l := W_{(t-1)N+l}$, one has $(\tilde{Z}_1, \ldots, \tilde{Z}_t) = g_{tN}(TN)(Z_1, \ldots, Z_t)$ hence indeed $\mathcal{G}_t = \sigma(Z_1, \ldots, Z_t)$ for $t = 1, \ldots, T$. $\square$

Chapter 4

PRELIMINARIES IN CONTINUOUS-TIME SETTING

1 Introduction

In this section we describe briefly part of the basic theory of the martingale problem and some results of weak convergence of laws in continuous time that are important to the subsequent development of our results. In chapter 3, we have seen how relevant the weak topology of probability measures is in order to obtain our results in the discrete-time setting.
The martingale problem is a celebrated approach to construct diffusion processes with continuous coefficients that was developed by Stroock D. and Varadhan S.R.S. in the seminal papers [54] and [55]. In the standard theory of stochastic differential equations, on the one hand, the probability space is fixed, and in order to construct an Itô process in this (arbitrary) probability space, one has to impose strong conditions on the coefficients, (for instance, locally Lipschitz and linear growth). The martingale problem overcomes these assumptions and ensures the existence of solutions, however in general, these solutions are not solutions in the strong sense, as those given by Itô's existence theorem. We will explain the concept of weak solutions and some notions concerning uniqueness. We follow [25] and [35].
As previously discussed in chapter 3, the functional in behavioural optimal investment depends on the law of the portfolio value processes, in continuous time the setting of weak solutions seems a natural one to adapt to our previous considerations.

2 Preliminaries

2.1 Canonical processes

Recall the definitions given in chapter 2, we denote by $\mathcal{W}_T^d$ the space of $\mathbb{R}^d-$ valued continuous functions i.e. $\mathcal{W}_T^d := C\left([0,T];\mathbb{R}^d\right)$.
We shall denote by $\mathcal{W}^d := C\left([0,\infty);\mathbb{R}^d\right)$, in this case the Borel $\sigma-$algebra is obtained by considering the topological space $\left(\mathcal{W}^d, \tau_{uc}\right)$ where τ_{uc} is the topology induced by the convergence

over compacts, or by the metric

$$d(f,g)=\sum_{n=1}^{\infty}\frac{1}{2^n}\frac{\|f-g\|_{[0,n]}}{1+\|f-g\|_{[0,n]}},$$

here the norm $\|\cdot\|$ denotes the uniform norm.

Definition 4.2.1. For each $t\in[0,T]$ denote by $x_t:\left(\mathcal{W}_T^d,\mathcal{B}\left(\mathcal{W}_T^d\right)\right)\to\left(\mathbb{R}^d,\mathcal{B}\left(\mathbb{R}^d\right)\right)$ the projection mappings, $x_t(w):=w_t$ for any continuous function $w\in\mathcal{W}_T^d$.
Define the projection on a finite collection $F=\{t_1,t_2,\dots t_n\}\subset[0,\infty)$ as the mapping $x_F:\left(\mathcal{W}_T^d,\mathcal{B}\left(\mathcal{W}_T^d\right)\right)\to\left(\mathbb{R}^{dn},\mathcal{B}\left(\mathbb{R}^{dn}\right)\right)$ given by $x_F(w):=(w_{t_i})_{t_i\in F}$.

Plainly, a canonical process is a $\mathbb{R}^d-$valued process on the measurable space $\left(\mathcal{W}_T^d,\mathcal{B}\left(\mathcal{W}_T^d\right)\right)$ whose value at time t is equal to $x_t(w.)=w(t)$ for $x\in\mathcal{W}_T^d$. Another important family of operators acting on $\mathcal{W}_T^d$ are defined below.

Definition 4.2.2. Let $t\in\mathbb{R}_+$, denote by $\alpha_t:\left(\mathcal{W}^d,\mathcal{B}\left(\mathcal{W}^d\right)\right)\to\left(\mathcal{W}^d,\mathcal{B}\left(\mathcal{W}^d\right)\right)$ we define the truncation operators as the mappings given by

$$\alpha_t(w.):=w_{.\wedge t}$$

The mappings α_t truncates the continuous function w up to its value at time t.
One can consider $\{x_t\}_{t\geqslant 0}$ as an stochastic process $x:[0,\infty)\times\mathcal{W}_T^d\to\mathbb{R}$, and notice that it is a continuous function, with respect the uniform norm.

Lemma 4.2.3. *The Borel $\sigma-$algebra is equal to the $\sigma-$algebra generated by projections $\mathcal{B}\left(\mathcal{W}^d\right)=\sigma\left(x_t:t\in\mathbb{R}_+\right)$*

Two probability measures $\mathbb{P}_1$, $\mathbb{P}_2$ on $\mathcal{B}\left(\mathcal{W}^d\right)$ are equal if and only if $\mathbb{P}_1\circ x_J^{-1}=\mathbb{P}_2\circ x_J^{-1}$ for each $J\subset\mathbb{R}_+$ finite.

2.2 Tightness and related results

Let $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$ be a complete filtered probability space, let d be an integer and consider an $\mathbb{R}^d$-valued stochastic process $\{X_t\}_{t\in[0,T]}$ having continuous trajectories, to each $t\in[0,T]$ the r.v. has a distribution $\mu_t(\cdot):=\mathbb{P}(X_t\in\cdot)$ i.e. a probability measure $\mu_t:\mathcal{B}\left(\mathbb{R}^d\right)\to[0,1]$. We refer to this measure as *the distribution at time t* or the *law of* X_t.
There is another probability measure associated to the process $\{X_t\}_{t\in[0,T]}$ if we regard the stochastic process $\{X_t\}_{t\in[0,T]}$ as a random variable that makes a correspondence between *each* ω and *the continuous function* $x.(\omega)$ (called a 'trajectory') then we have a $\mathcal{W}_T^d$ valued r.v. and we can consider the "distribution" $Q:\mathcal{B}\left(\mathcal{W}_T^d\right)\to[0,1]$ of this random variable. In this case, we shall say that Q is *the law of* the process $\{X_t\}_{t\in[0,T]}$. This means that if $\Gamma\in\mathcal{B}\left(\mathcal{W}_T^d\right)$ then

$$Q(\Gamma):=\mathbb{P}(\omega:X.(\omega)\in\Gamma).$$

Formally, fix an stochastic process $X(t,\omega)$ and define the mapping $\varphi_x:(\Omega,\mathcal{F})\to\left(\mathcal{W}^d,\mathcal{B}\left(\mathcal{W}^d\right)\right)$

$$\varphi_x(\omega):=X.(\omega)$$

It is clear that the mapping φ_x is $\mathcal{B}\left(\mathcal{W}^d\right)\backslash\mathcal{F}-$ measurable, taking a set in the generating class of the σ- algebra $\mathcal{B}\left(\mathcal{W}^d\right)$, if $F=\{t_1,t_2,\dots,t_n\}$ and $\Gamma_F:=\left\{w\in\mathcal{W}^d:w_{t_1}\in A_1,\dots w_{t_n}\in A_n\right\}$

with $A_i \in \mathcal{B}\left(\mathbb{R}^d\right)$ then it is easy to see that

$$\varphi_x^{-1}\left(\Gamma_F\right) = \left\{\omega \,:\, X_{t_i}\left(\omega\right) \in A_i \,, t_i \in A_i\right\} \in \mathcal{F}.$$

Then the law of $\{X_t\}_{t\in[0,T]}$ is given by the distribution of φ_x i.e. $\mathbb{P} \circ \varphi_x^{-1}$.

$$Q\left(\Gamma\right) = \mathbb{P}\left(\omega \,:\, \varphi_x\left(\omega\right) \in \Gamma\right), \quad \text{for } \Gamma \in \mathcal{B}\left(\mathcal{W}^d\right). \tag{4.2.1}$$

The same holds for the space of functions with cádlág trajectories, but the details in this case are beyond the scope of this work. We will be considering processes having continuous sample paths.
As it was explained in chapter 2, if $(E, \mathcal{E})$ is a measurable space, we define $\mathcal{M}^1\left(E\right)$ the set of probability measures on $\mathcal{E}$. Recall the following definitions.

Definition 4.2.4. Suppose E is a Polish space and $\mathcal{E} = \mathcal{B}\left(E\right)$ its Borel $\sigma-$algebra. Consider a collection of probability measures $\mathcal{F} \subseteq \mathcal{M}^1\left(E\right)$, we say that $\mathcal{F}$ is *tight* if for any $\epsilon > 0$ there is a compact $K_\epsilon \subset E$ such that

$$\sup_{\mu\in\mathcal{F}} \mu\left(E \backslash K_\epsilon\right) < \epsilon. \tag{4.2.2}$$

As mentioned in chapter 2, $\mathcal{M}^1\left(E\right)$ with the topology induced by weak convergence is a Polish space.

Definition 4.2.5. Let $C_b\left(E\right)$ be the space of real-valued, bounded functions on E a Polish space, we say that a sequence $\{\mu_n\}_{n\geqslant 1} \subset M^1\left(E\right)$ converges weakly to the probability measure μ if for any $f \in C_b\left(E\right)$

$$\lim_{n\to\infty} \int_E f\left(x\right) \mu_n\left(dx\right) = \int_E f\left(x\right) \mu\left(dx\right).$$

And we write $\mu_n \Rightarrow \mu$ and say that μ_n converges weakly to μ or that μ_n is a weakly converging sequence.

We state without proof a characterisation of tightness of a set of laws of continuous processes. As a matter of fact, this is a consequence of the Arzelá-Ascoli theorem.

Proposition 4.2.6. *Suppose $\{\mathbb{P}_\lambda\}_{\lambda\in\Lambda}$ is the set of laws of the processes $\left\{X_t^\lambda \,:\, t \in [0,T]\right\}_{\lambda\in\Lambda}$ with continuous trajectories, suppose that for any $\epsilon > 0$ there is $\alpha > 0$ such that*

$$\sup_{\lambda\in\Lambda} \mathbb{P}_\lambda\left(x \in \mathcal{W}^d \,:\, \left|x\left(0\right)\right| > \alpha\right) < \epsilon, \tag{4.2.3}$$

And, suppose there exist constants $K_T > 0$ and $\alpha, \beta > 0$ such that

$$\sup_{\lambda\in\Lambda} \mathbf{E}\left|X_t^{(\lambda)} - X_s^{(\lambda)}\right|^\beta \leqslant K_T \left|t-s\right|^{1+\alpha} \ \textit{for all } s,t \in \left[0,T\right], \tag{4.2.4}$$

then the family of laws, $\{\mathbb{P}_\lambda\}_{\lambda\in\Lambda}$ with $\mathbb{P}_\lambda := \mathbb{P} \circ \left(X^{(\lambda)}\right)^{-1}$ is tight.

2.3 The martingale problem and weak solutions of SDE's

In this section we will describe the concept of weak solution of a stochastic differential equation and the martingale problem formulation. The martingale approach was developed by Stroock

D. and Varadhan S.R.S. in [54] and [55].
This approach will allow us to analyse the problem of optimal investment under behavioural criteria, it remains an open question whether there is an approach to investigate optimality on a fixed probability space and without making stringent assumptions on the model. We follow closely [25] and [56].

As before, consider the topological space $(\mathcal{W}^d, \tau_{uc})$ with τ_{uc} the topology induced by the uniform convergence of continuous functions over compacts and consider the Borel σ-field induced by the collection of the $\tau_{uc}-$ open sets, denoted by $\mathcal{B}(\mathcal{W}^d)$ and $\mathcal{N}_t \subset \mathcal{B}(\mathcal{W}^d)$ the sub-σ-fields generated by the mappings $\alpha_s : \mathcal{W}^d \to \mathcal{W}^d$ for $s \leqslant t$, in other words, $\mathcal{N}_t := \sigma(\{\alpha_s : s \leqslant t\})$ see Definition 4.2.2.

We define a class of 'path functionals' whose elements are the main components of a stochastic differential equation, these functionals are generally called 'the coefficients' of a stochastic Itô equation.

Definition 4.2.7. We denote by $\mathcal{A}^{d,r}$ the set of all functions $\alpha : [0,\infty) \times \mathcal{W}^d \to \mathbb{R}^{d\times r}$ such that

1. it is $\mathcal{B}([0,\infty)) \times \mathcal{B}(\mathcal{W}^d) / \mathcal{B}(\mathbb{R}^{d\times r})$ measurable, and
2. for each $t \in [0,\infty)$, the mapping $w \to \alpha(t,w)$, is $\mathcal{N}_t/\mathcal{B}(\mathbb{R}^{d\times r})$- measurable.

Where the Borel $\sigma-$field $\mathcal{B}(\mathbb{R}^{d\times r})$ is obtained by identifying the space of real-valued matrices $\mathbb{R}^{d\times r}$ with the vector space $\mathbb{R}^{dr}$ and its Borel $\sigma-$algebra is given by $\mathcal{B}(\mathbb{R}^{dr})$.
Given $\alpha \in \mathcal{A}^{d,r}$ and $\beta \in \mathcal{A}^{d,1}$. Consider the $d-$dimensional stochastic differential equation and a solution of such an equation, $X = \{X_t\}_{t\geqslant 0}$

$$dX_t^i = \sum_{j=1}^{r} \alpha^{ij}(t, X.)\, dB^j(t) + \beta^i(t, X.)\, dt, \qquad i = 1, 2, \ldots, d. \tag{4.2.5}$$

Definition 4.2.8. Following [25], we define a solution of the equation (4.2.5) as a $d-$dimensional continuous process X defined *on a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ such that

- There exists an r-dimensional $\mathcal{F}_t-$Wiener process $B = \{B(t)\}_{t\geqslant 0}$ with $B(0) = 0$.
- The $d-$dimensional continuous process $X = \{X(t)\}_{t\geqslant 0}$ is adapted to $\{\mathcal{F}_t\}_{t\geqslant 0}$ i.e. X is a mapping such that, for each $t \in [0,\infty)$, $X(t,\cdot)$ is $\mathcal{F}_t/\mathcal{B}_t(\mathcal{W}^d)$ $-$measurable.
- the set of adapted processes $\Phi^{ij}(t,\omega)$ and $\Psi^i(t,\omega)$, usually called the *coefficients* of the equation and are defined by

$$\Phi^{ij}(t,\omega) = \alpha^{ij}(t, X.(\omega)), \tag{4.2.6}$$

$$\Psi^i(t,\omega) = \beta^i(t, X.(\omega)), \tag{4.2.7}$$

  are measurable, $\mathcal{F}_t-$adapted and such that Φ^{ij} are locally square integrable and Ψ^i locally integrable[1] respectively, $\mathbb{P} - a.s.$

[1] A process is locally integrable if it is an $\mathcal{F}_t-$adapted process and for every $t \geqslant 0$, $\int_0^t |f_s|\, ds < \infty$ $a.s.$

- with probability one $X(t) = \left(X^1(t), \ldots, X^d(t)\right)$ and $B(t) = \left(B^1(t), \ldots, B^r(t)\right)$ satisfy

$$X^i(t) - X^i(0) = \sum_{j=1}^{r} \int_0^t \alpha^{ij}(s, X.)\, dB^j(s) + \int_0^t \beta^i(s, X.)\, ds. \tag{4.2.8}$$

As a matter of fact one can regard the pair (X, B) as a solution to (4.2.5) or say that $X(t)$ is a solution with the Wiener process $B = \{B(t)\}_{t \geqslant 0}$.
Saying that (X, B) solves (4.2.8) is equivalent to

$$\mathbb{P}\Bigg(X^i(t) - X^i(0) = \sum_{j=1}^{r} \int_0^t \alpha^{ij}(s, X)\, dB^j(s) + \int_0^t \beta^i(s, X)\, ds, \text{ for } i = 1, 2 \ldots, d,\ t \geqslant 0\Bigg) = 1.$$

Remark 4.2.9. Definition 4.2.8 differs from the usual concept of "strong" solution of an SDE, first of all, the probability space and the Wiener process $B(t)$ are not fixed beforehand. Secondly, the process $X(t)$ is not adapted to the natural filtration of the Wiener process $B(t)$. Thus, one may even consider the set $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0}, \mathbb{P}, \{B(t)\}_{t \geqslant 0}, \{X(t)\}_{t \geqslant 0}\right)$ as a solution and in this case the most important component of a solution is the law of the process X.

The solution described in Definition 4.2.8 and in Remark 4.2.9 is generally called a *weak solution*.
On the other hand when investigating the concept of strong solution and the conditions on which such solutions exist, an important concept of uniqueness that emerges is the *pathwise uniqueness*.

Definition 4.2.10. Pathwise uniqueness holds if given two solutions X and X' of (4.2.8) *on the same filtered probability space* $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0}, \mathbb{P}\right)$ and with respect the same Wiener processes $B(t)$ such that $X(0) = X'(0)$ a.s then

$$\mathbb{P}\left(\omega : X(t, \omega) = X'(t, \omega) \text{ for all } t\right) = 1.$$

When considering two weak solutions of an equation, as the solutions may not be defined on the same probability space, the concept of pathwise uniqueness in meaningless, however it is possible to "compare" the laws of the processes (as these are measures on the same space $\mathcal{W}^d$). Thus, we have the following central concept of uniqueness.

Definition 4.2.11. We say that *uniqueness in law* holds if whenever X and X' are two solutions such that $X(0)$ and $X'(0)$ have the same law (in $\mathbb{R}^d$) then the processes X and X' have the same law in $\mathcal{B}\left(\mathcal{W}^d\right)$.

The usual setting of stochastic differential equations are a particular case of (4.2.5).

Definition 4.2.12. Let $\sigma(t, x) = \left(\sigma^{ij}(t, x)\right)_{ij}$ be a Borel measurable function $\sigma : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^{d \times r}$ and $b(t, x) = \left(b^i(t, x)\right)_{i \leqslant d}$ is a Borel measurable function $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$. Then we define functionals Ψ and Φ in (4.2.6) by $\alpha(t, w) := \sigma(t, w(t))$ and $\beta(t, w) := b(t, w(t))$. In such case the stochastic differential equation is said to be of *Markovian type*.

Consider the equation (4.2.5) with $\alpha \in \mathcal{A}^{d,r}$ and $\beta \in \mathcal{A}^{d,1}$. For $f \in C_b^2\left(\mathbb{R}^d; \mathbb{R}\right)$ let us define the following operator

$$(\mathcal{L}f)(t, w) := \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(t, w) \frac{\partial^2 f}{\partial x^i \partial x^j}(w(t)) + \sum_{i=1}^{d} \beta^i(t, w) \frac{\partial f}{\partial x^i}(w(t)), \quad t \in [0, \infty), w \in \mathcal{W}^d, \tag{4.2.9}$$

where

$$a^{ij}(t,w) = \sum_{k=1}^{r} \alpha^{ik}(t,w)\,\alpha^{jk}(t,w). \tag{4.2.10}$$

We shall denote $\langle \cdot,\cdot \rangle$ the inner product in $\mathbb{R}^d$. If $\{N_t,\mathcal{F}_t\}_{t\geqslant 0}$ is a local martingale we denote by $\langle N.,N.\rangle_t$ the Doob-Meyer process, it is the unique continuous process such that $\left\{N_t^2 - \langle N,N\rangle_t, \mathcal{F}_t\right\}$ is a local martingale. From our notation and the context, we deem that there is no ambiguity. We have the following remark, this is a consequence of Itô's formula.

Remark 4.2.13. Suppose (X,B) solves (4.2.5) on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and with respect some filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ the process $B = \left\{B(t)\right\}_{t\geqslant 0}$ is an $\mathcal{F}_t-$Wiener process. Then we have the following

i) We have for all $t \geqslant 0$

$$X(t) - X(0) - \int_0^t \beta(s,X)\,ds = \int_0^t \alpha(s,X)\,dB(s), \tag{4.2.11}$$

is a local martingale, in other words, each coordinate $M^i(t) := X^i(t) - X^i(0) - \int_0^t \beta^i(s,X)\,ds$ is a real valued local martingale w.r.t to $\left(\{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$.

ii) Moreover, there is $a : [0,\infty) \times \Omega \to \mathbb{R}^{d\times r}$, given by (4.2.10), it is locally integrable, a symmetric, non-negative matrix for all (s,w) and $a \in \mathcal{A}^{d,d}$ such that

$$\left\langle M^i, M^j \right\rangle_t = \int_0^t a_s^{ij}\,ds. \tag{4.2.12}$$

iii) For any $f \in C^2([0,T];\mathbb{R})$,

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}f(s,X.)\,ds \in \mathcal{M}^c_{loc}.^2 \tag{4.2.13}$$

The following definition and part of the presentation given below are based on the ideas of chapter III in [35], definition III.9.1.

Definition 4.2.14. Suppose M_t is a $\mathbb{R}^d$-valued local martingale and $a : \Omega \times [0,\infty) \to S^d_+$ i.e. $a(t,\omega)$ is no-negative symmetric matrix. We say that a local martingale $\left\{M_t,\mathcal{F}_t\right\}_{t\geqslant 0}$ is *admissible* local martingale if (4.2.12) holds.

As sometimes it is clear that the martingale property is with respect to the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$, we omit its inclusion.

Suppose $X = \{X(t)\}_{t\geqslant 0}$ is a $d-$dimensional continuous adapted process defined on a complete filtered probability space $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathbb{P}\right)$, such that (4.2.13) holds for any $f \in C^2\left([0,T];\mathbb{R}^d\right)$. We claim that such a condition implies that the process in the left hand side of (4.2.11) is an admissible local martingale, and $a(s,X) \in \mathcal{A}^{d,r}$ is the process such that (4.2.12) holds.

Indeed, let $B_r = \left\{x \in \mathbb{R}^d \mid |x| \leqslant r\right\}$, choosing $f \in C_b^2\left(\mathbb{R}^d\right)$ such that $f(x) = x_i$ for $x \in B_r$, setting $\sigma_r = \inf\{t : X(t) \notin B_r\}$. Define for $l = 1,2\ldots$ we have

$$M_l^i(t) = X^i(t\wedge\sigma_l) - X^i(0) - \int_0^{t\wedge\sigma_l} \beta^i(s,X)\,ds \in \mathcal{M}^{2,c\ 3}, \quad i = 1,2,\ldots,d, \tag{4.2.14}$$

[2] Here, $\mathcal{M}^c_{loc}$ denotes the space of continuous local martingales.
[3] Here, $\mathcal{M}^{2,c}$ denotes the space of continuous square integrable martingales.

by a direct application of (4.2.13). Define the processes with $y \in \mathbb{R}^d$

$$M_t^y := \left\langle y, X(t) - X(0) - \int_0^t \beta(s,X)ds \right\rangle$$

for each $y \in \mathbb{R}^d$, it is a continuous local martingale as it is the sum of local martingales multiplied by constants $y_i \in \mathbb{R}$. In fact, its Doob-Meyer process is given by the increasing process

$$\langle M^y, M^y \rangle_t = \int_0^t \langle y, a(s,X)\, y \rangle\, ds. \tag{4.2.15}$$

To see this, first, notice that we can assume $X(0) = 0$ as $\langle y, X(0)\rangle$ is a martingale. By assumption, X satisfies the condition (4.2.13), choosing $f(x) = \langle y, x\rangle^2$ on B_l we have

$$\langle y, X(t)\rangle^2 - 2\int_0^t \langle y, \beta(s,X)\rangle \langle y, X(s)\rangle\, ds - \int_0^t \langle y, a(s,X)\, y\rangle\, ds, \tag{4.2.16}$$

is a local martingale.
Denote by $N_t^y = \langle y, X_t\rangle^2 - 2\int_0^t \langle y, \beta(s,X)\rangle\langle y, X_s\rangle\, ds$ it is enough to prove that $(M_t^y)^2 - N_t^y$ is a local martingale as this would imply that

$$(M_t^y)^2 - \int_0^t \langle y, a(s,X)\, y\rangle\, ds - \left(N_t^y - \int_0^t \langle y, a(s,X)\, y\rangle\, ds\right)$$

is a local martingale, thus the first difference would be a local martingale as well. On the other hand,

$$(M_t^y)^2 = \left\langle y, X(t) - \int_0^t \beta(s,X)\, ds\right\rangle = \langle y, X(t)\rangle M_t^y - \left\langle y, M_t^y \left(\int_0^t \beta(s,X)\, ds\right)\right\rangle =$$

$$= \langle y, X(t)\rangle^2 - \langle y, X(t)\rangle \left\langle y, \int_0^t \beta(r,X)\, dr\right\rangle ds - 2\int_0^t \int_0^s \langle y, \beta(r,X)\rangle\langle y, \beta(s,X)\rangle\, drds.$$

The process N_t^y can be written in the following form

$$N_t^y = \langle y, X(t)\rangle^2 - 2\int_0^t \langle y, \beta(s,X)\rangle M_s^y ds - 2\int_0^t \int_0^s \langle y, \beta(s,X)\rangle\langle y, \beta(r,X)\rangle\, drds.$$

Then

$$(M_t^y)^2 - N_t^y = 2\int_0^t \langle y, \beta(s,X)\rangle M_s^y ds - \int_0^t \langle y, \beta(s,X)\rangle M_t^y ds + 2\int_0^t \int_0^s \langle y, \beta(s,X)\rangle\langle y, \beta(r,X)\rangle\, drds$$

$$- \langle y, X(t)\rangle \int_0^t \langle y, \beta(s,X)\rangle\, ds.$$

The last integral can be expressed in terms of M_t^y, indeed $\langle y, X(t)\rangle \int_0^t \langle y, \beta(s,X)\rangle\, ds$ is equal to

$$\left\langle y, X_t - \int_0^t \beta(s,X)\, ds\right\rangle \int_0^t \langle y, \beta(s,X)\rangle\, ds + \left[\int_0^t \langle y, \beta(s,X)\rangle\, ds\right]^2,$$

therefore

$$(M_t^y)^2 - N_t^y = 2\int_0^t \langle y, \beta(s,X)\rangle (M_s^y - M_t^y)\, ds + 2\int_0^t \int_0^s \langle y, \beta(s,X)\rangle\langle y, \beta(r,X)\rangle\, drds - \ldots$$

$$-\left[\int_0^t \langle y, \beta(s,X)\rangle\, ds\right]^2,$$

then, applying Fubini's theorem on the last integrals (involving $\beta(s,X)$, $\beta(r,X)$)

$$(M_t^y)^2 - N_t^y = 2\int_0^t (M_s^y - M_t^y)\, dA(s),$$

where $A_t^y = \int_0^t \langle y, \beta(s,X)\rangle ds$ is of finite variation, then applying the integration by parts formula

$$(M_t^y)^2 - N_t^y = 2\int_0^t (M_s^y - M_t^y)\, dA_s = -2\left(M_t^y A_t^y - \int_0^t M_s^y dA_s^y\right) = -2\int_0^t A_s^y dM_s^y.$$

And the last expression is a local martingale, then $N_t^y - \int_0^t \langle y, a(s,X)\, y\rangle ds$ is a local martingale, then (4.2.13) implies that $\{M_t\}_{t\geqslant 0}$ is a local martingale and $\{M_t^y\}_{t\geqslant 0}$ is a local martingale with Doob-Meyer process equal to $\int_0^t \langle y, a(s,X)\, y\rangle ds$.

Remark 4.2.15. In particular, taking $y = e_i = (\delta^{ij})_{j=1,\ldots,n}$ we obtain that

$$\langle M^i, M^i\rangle_t = \int_0^t a^{ii}(s,X)\, ds.$$

And if, $y = e_i + e_j$ it follows that

$$\langle M^i + M^j, M^i + M^j\rangle_t = \int_0^t a^{ii}(s,X) + 2a^{ij}(s,X) + a^{jj}(s,X)\, ds$$

Thus

$$\langle M^i, M^j\rangle_t = \int_0^t a^{ij}(s,X)\, ds, \tag{4.2.17}$$

this means M_t is an *admissible* $d-$dimensional local martingale in the of sense Definition 4.2.14.

We now see that, as matter of fact, condition (4.2.13) is equivalent to having $M_t^y = \langle y, M_t\rangle$ to be an admissible local martingale with (4.2.15). Indeed, if $M_t = X(t) - X(0) - \int_0^t \beta(s,X)\, ds$ is an admissible local martingale the so-called "martingale version" of Itô's formula can be applied without knowing that M_t^i is an Itô process, whatsoever. The theorem below is taken from [35], Theorem III.9.5.

Theorem 4.2.16. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space, with $\mathcal{F}_0$, $\mathbb{P}$ complete. Let $b : \Omega\times[0,\infty) \to \mathbb{R}^d$, be an $\mathcal{F}_t-$adapted and $\mathcal{B}([0,\infty))\otimes\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable process, such that*

$$\int_0^t |b(s)|\, ds < \infty \quad \textit{for all } t \geqslant 0.$$

Define

$$\eta_t := \int_0^t b(s)\, ds.$$

Consider the continuous semimartingale $\zeta_t := \xi_t + \eta_t$. Where ξ_t is an admissible local martingale with a_t and ξ_t satisfying (4.2.12). Let $u(x)$ be a real-valued function having continuous derivatives $u_{x_i}, u_{x_i x_j}$ on $\mathbb{R}^d$ for all $i,j = 1,\ldots,d$. Then the stochastic process

$$u(\zeta_t) - \int_0^t \mathcal{L}_s^{a,b} u(\zeta_s)\, ds, \tag{4.2.18}$$

where $\mathcal{L}_s^{a,b}$ is the differential operator given by

$$\mathcal{L}_s^{a,b} u(x) = \sum_{i=1}^{d} b_i(s)\, u_{x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(t)\, u_{x_i x_j}(x), \tag{4.2.19}$$

is a local martingale w.r.t $\left(\{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$.

From the remarks given above and Theorem 4.2.16 we have the following corollary

Corollary 4.2.17. *Let* $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ *be a complete filtered probability space, let* $\{X_t\}_{t\geqslant 0}$ *be an* $\mathbb{R}^d$*- valued continuous process.*
Then the following conditions are equivalent.

The process $M_t = X_t - X_0 - \int_0^t \beta(s, X)\, ds$, *and the real valued processes* $\langle y, M_t\rangle$ *are local martingales w.r.t to* $\left(\{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$, *for any* $y \in \mathbb{R}^d$, *further each local martingale* M_t^y *has a Doob-Meyer process that is equal to*

$$\langle\langle y, M.\rangle\rangle_t = \langle M^y_\cdot\rangle_t = \int_0^t \langle y, a(s, X)\, y\rangle\, ds \tag{4.2.20}$$

In other words, M_t *and* $M_t^y = \langle y, M_t\rangle$ *are admissible local martingales w.r.t* $\left(\mathcal{F}_t, \mathbb{P}\right)$.

For any $u \in C_0^\infty\left(\mathbb{R}^d; \mathbb{R}\right)$ *the process*[4]

$$u(X_t) - u(X_0) - \int_0^t \mathcal{L}_s u(X_s)\, ds \tag{4.2.21}$$

is a local martingale w.r.t. $(\mathcal{F}_t, \mathbb{P})$.

These remarks and connections between different families of martingales associated to a process X solving (4.2.5) motivate the following crucial definition, namely, the concept of a *martingale problem* and its solution.

Definition 4.2.18. Consider the canonical process, $\{x_t\}_{t\geqslant 0}$ on $\left(\mathcal{W}^d, \mathcal{B}\left(\mathcal{W}^d\right)\right)$, and the filtration given by $\mathcal{F}_t = \mathcal{N}_t$. Let $\mu \in \mathcal{M}^1\left(\mathbb{R}^d\right)$ and $(s, x) \in [0, \infty) \times \mathbb{R}^d$ and $\mathcal{L}_t$ the operator given by (4.2.9) (or (4.2.19)), we say that a probability measure on $\mathcal{B}\left(\mathcal{W}^d\right)$, $\mathbb{P}_{(s,\mu)}$ is a *solution to the martingale problem* $(s, \mu, \mathcal{L}_t)$ if

- The condition (4.2.21) holds for x_t for all $t \geqslant s$ i.e.

$$u(x_t) - u(x_s) - \int_s^t \mathcal{L}_r u(x_r)\, dr \text{ is an } \left(\mathcal{F}_t, \mathbb{P}_{(s,\mu)}\right) \text{ local martingale,} \tag{4.2.22}$$

  for all $u \in C^2\left(\mathbb{R}^d\right)$.

- For all $\Gamma \in \mathcal{B}\left(\mathbb{R}^d\right)$ we have $P_{s,\mu}\left(w \in \mathcal{W}^d : x_t(w) \in \Gamma \text{ for } 0 \leqslant t \leqslant s\right) = \mu(\Gamma)$.

Let $s \geqslant 0$ and $x \in \mathbb{R}^d$, we say $(s, x, \mathcal{L}_t)$ is a martingale problem with constant initial condition at time s, i.e. $x(s) = x$, in other words, the martingale problem $(s, \varepsilon_x, \mathcal{L}_t)$, with ε_x is the Dirac's measure. Denote by $P_{s,x}$ a solution of the martingale problem $(s, x, \mathcal{L}_t)$.

Example 4.2.19. The solution to the martingale problem w.r.t. to the triplet $\left(0, \varepsilon_0, \frac{1}{2}\Delta^d\right)$ where Δ^d is the Laplacian operator in $\mathbb{R}^d$ is called the *Wiener measure*.

[4] We denote by $C_0^\infty\left(\mathbb{R}^d; \mathbb{R}\right)$ the space of smooth functions with compact support.

Given $a \in \mathcal{A}^{d,d}$ and $b \in \mathcal{A}^{d,1}$ it is possible to reformulate the martingale problem (Definition 4.2.18) in alternative ways, this is the content of the following proposition.

Proposition 4.2.20. *Let* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *be a complete filtered probability space, suppose* $\{X(t)\}_{t \geqslant s}$ *is a continuous and* $\mathcal{F}_t-$*adapted stochastic processes and let* $\alpha \in \mathcal{A}^{d,r}$, $\beta \in \mathcal{A}^{d,1}$ *and* $a : [s, \infty) \times \Omega \to S_+^d$ *given by (4.2.10) such that* $X(s) = x$.
Then, the following conditions are equivalent

I *For each* $f \in C_0^\infty(\mathbb{R}^d)$ *the process* $\left\{C_t^{s,f}\right\}_{t \geqslant s}$ *defined by*

$$C_t^{s,f} := f(X(t)) - f(X(s)) - \int_s^t \mathcal{L}_r f(X(r))\, dr, \tag{4.2.23}$$

is an $\mathcal{F}_t-$*martingale.*

II *For each* $f \in C_0^2(\mathbb{R}^d)$ *the process* $\left\{C_t^{s,f}\right\}_{t \geqslant s}$ *defined by (4.2.23) is an* $\mathcal{F}_t-$*martingale.*

III *For each* $f \in C^2(\mathbb{R}_+ \times \mathbb{R}^d)$ *the process* $\left\{C_t^{s,f}\right\}_{t \geqslant s}$ *defined by*

$$C_t^{s,f} = f(t, X(t)) - f(s, X(s)) - \int_s^t \mathcal{L}_r f(r, X(r))\, dr,$$

is an $\mathcal{F}_t-$*local martingale.*

IV *For each* $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d; \mathbb{R})$, *i.e. bounded,* $C^{1,2}$ *functions, the process* $\left\{\mathcal{C}_t^{s,f}\right\}_{t \geqslant s}$ *defined by*

$$\mathcal{C}_t^{s,f} = f(t, X(t)) - f(s, X(s)) - \int_s^t \mathcal{L}_u f(u, X(u))\, du,$$

is an $\mathcal{F}_t-$*martingale.*

V *For each* $\theta \in \mathbb{R}^d$ *the process* $\{\chi_\theta^s(t)\}_{t \geqslant s}$ *defined by*

$$\chi_\theta^s(t) = \exp\left(\left\langle \theta, X(t) - X(s) - \int_0^t \beta(s, X)\, ds \right\rangle - \frac{1}{2}\int_0^t \langle \theta, a(s, X)\, \theta \rangle\, ds\right), \tag{4.2.24}$$

is an $\mathcal{F}_t-$*martingale.*

VI *For each* $\theta \in \mathbb{R}^d$ *the process* $\{\chi_{i\theta}^s(t)\}_{t \geqslant s}$ *defined by*

$$\chi_{i\theta}^s(t) = \exp\left(i\left\langle \theta, X(t) - X(s) - \int_0^t \beta(s, X)\, ds \right\rangle + \frac{1}{2}\int_0^t \langle \theta, a(s, X)\, \theta \rangle\, ds\right), \tag{4.2.25}$$

is an $\mathcal{F}_t-$*local martingale.*

For a complete proof of this proposition see [53], [56] or [52].

Results on solutions of the martingale problem

An important and fairly general result on the existence of solutions to the martingale problem is the following.

Theorem 4.2.21. *Suppose the functionals a and b belong to $\mathcal{A}^{d,d}$ and $\mathcal{A}^{r,1}$ respectively. Further, suppose that $a : [s, \infty) \times \mathcal{W}^d \to S^d_+$,[5] and $b : [s, \infty) \times \mathcal{W}^d \to \mathbb{R}^d$ are continuous and bounded. Then, there exist a solution $P_{s,x}$ to the martingale problem $(\mathcal{L}_t, s, x)$ for any $s \in \mathbb{R}_+$ and $x \in \mathbb{R}^d$.*

For a proof of this fact, refer to [25], Theorem 4.3.3.

Reduction of the martingale problem to time-homogeneous coefficients.

In the case of having a differential operator whose coefficients a and b depend on the state of the process rather that on the trajectories, it is possible to formulate the martingale problem for time-dependent coefficients in terms of a time-homogeneous martingale problem.
Suppose $a : [0, \infty) \times \mathbb{R}^d \to S^d_+$ and $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ are bounded and measurable functions and let $\mathcal{L}_t$ be the differential operator associated to a, b see (4.2.19) on $C^2_b\left(\mathbb{R}^d; \mathbb{R}\right)$.
Denote any element $\tilde{x} \in \mathbb{R}^{d+1}$ by $\tilde{x} = (x_0, x)$ with $x_0 \in \mathbb{R}$ and $x \in \mathbb{R}^d$ and define the 'time-homogeneous' coefficients $\tilde{a}$ and $\tilde{b}$ by

$$\tilde{a}^{ij}(\tilde{x}) = \begin{cases} 0, & \text{if } i \text{ or } j = 0, \\ a^{ij}(x_0 \vee 0, x) & \text{otherwise,} \end{cases} \tag{4.2.26}$$

$$\tilde{b}^{i}(\tilde{x}) = \begin{cases} 1, & \text{if } i = 0, \\ b^{i}(x_0 \vee 0, x) & \text{otherwise.} \end{cases} \tag{4.2.27}$$

These coefficients define a differential operator $\widetilde{\mathcal{L}} = \sum_{i,j=0}^d \tilde{a}^{ij}(\tilde{x}) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=0}^d b^i(\tilde{x}) \frac{\partial}{\partial x_i}$.
It is possible to check that there is a one to one correspondence between solutions of the martingale problem $(\mathcal{L}_t, s, \epsilon_x)$ and $\left(\widetilde{\mathcal{L}}, 0, \epsilon_{\tilde{x}}\right)$ with $\tilde{x} = (s, x)$. Define the following mapping $\Psi_s(\omega) : \mathcal{W}^d \to \mathcal{W}^{d+1}$ to be $\Psi_s(\omega) = (\cdot + s, \omega(\cdot + s)) \in \mathcal{W}^{d+1}$ and if $\tilde{x}$ denotes the canonical process on $\mathcal{W}^{d+1}$ we have $\tilde{x}(t, \Psi_s(\omega)) = (t + s, x(t + s, \omega))$. This means Ψ_s "shifts" the path to the time s and embeds the time coordinate as another 'constant function'.

Remark 4.2.22. The probability measure $P \in \mathcal{M}^1\left(\mathcal{W}^d\right)$ solves the martingale problem associated to $(\mathcal{L}_t, s, \epsilon_x)$ if and only if $P \circ \Psi_s^{-1} \in \mathcal{M}^1\left(\mathcal{W}^{d+1}\right)$ solves the 'time-homogeneous' martingale problem $\left(\widetilde{\mathcal{L}}, 0, \epsilon_{\tilde{x}}\right)$.

Indeed, by Proposition 4.2.22, P is a solution to the martingale problem $(\mathcal{L}_t, s, x)$ if and only if, for any $f \in C^{1,2}_b\left(\mathbb{R}^+ \times \mathbb{R}^d; \mathbb{R}\right)$ and for any $t > s \geqslant 0$,

$$f(t+s, x_{t+s}(\omega)) - f(s, x_s(\omega)) - \int_s^t \left(\frac{\partial}{\partial x_i} + \mathcal{L}_r\right) f(r, x_r(\omega))\, dr \text{ is a martingale w.r.t } \left(\mathcal{N}_t, P\right),$$

making a change of variables and using the definition of the mapping Ψ_s it follows that

$$\int_s^{s+t} \frac{\partial}{\partial u} f(u, x_u)\, du = \int_0^t \frac{\partial}{\partial u} f(\tilde{x}(u, \Psi_s(\omega)))\, du = \int_0^t b^0(\tilde{x}(u, \Psi_s(\omega))) \frac{\partial}{\partial x_0} f(\tilde{x}(u, \Psi_s(\omega)))\, du,$$

[5] Here $S^d_+ \subset \mathbb{R}^{d \times d}$ denotes the set of positive semidefinite symmetric matrices

similarly

$$\int_s^{t+s} \mathcal{L}_r f(r, x_r(\omega))\, dr = \int_0^t \mathcal{L} f(r+s, x_{r+s}(\omega))\, dr = \tag{4.2.28}$$

$$\int_0^t \mathcal{L} f(\tilde{x}_r(\Psi_r(\omega)))\, dr, \tag{4.2.29}$$

we obtain that for any $t > r \geqslant s$ and $A \in \mathcal{N}_r$,

$$\int_A f(\tilde{x}_t(\Psi_s(\omega))) - f(\tilde{x}_r(\Psi_s(\omega))) - \int_s^t \tilde{\mathcal{L}} f(\tilde{x}_u(\Psi_s(\omega)))\, du dP = 0. \tag{4.2.30}$$

And using the formula connecting the measure $P \circ \Psi_s^{-1}$ and the integrals of the form $\int_{\mathcal{W}^d} g(\Psi_s(\omega))\, dP$ as in (4.2.30) (functions of Ψ_s) we have that (4.2.30) is equal to

$$\int_A f(\tilde{x}(t, \omega)) - f(\tilde{x}(r, \omega)) - \int_r^t \tilde{\mathcal{L}} f(\tilde{x}(u, \omega))\, du d(P \circ \Psi_s) = 0. \tag{4.2.31}$$

This means that $P \circ \Psi_s^{-1}$ is a solution of the time-homogeneous martingale problem $(\tilde{\mathcal{L}}, 0, \epsilon_{\tilde{x}})$. And the equivalence follows.

Conversely, one can prove that, if the mapping Φ_s shifts the trajectories from time s onwards to the origin, i.e. $\Phi_s : \Omega^{d+1} \to \Omega^d$, with the mapping Φ_s being such that $\tilde{x}(t, \Phi_s(\tilde{\omega}))$ then $\tilde{P}$ is a solution for the martingale problem in for $(\mathcal{L}_t, s, x)$ if and only if $\tilde{P} \circ \Psi_s$ is a solution to the martingale problem $(\tilde{\mathcal{L}}, 0, \tilde{x})$.
Using this reduction we can solve 'inhomogeneous' martingale problems using results on time-homogeneous martingale problems.

Martingale problems and weak solutions

How is the martingale problem related to the concept of a weak solution?
The following lemmata allow to link the concepts of weak solution of (4.2.5) and a solution to the martingale problem. Consider the case when a and β depend on the state, i.e. $a(t, w) = a(t, w(t))$.
First of all, notice that the solution of a martingale problem is given by a family of measures $\{P_{s,x}\}$ (as s and x varies). An approach explaining such a link is given by the following remark: there is an *extension* of the stochastic basis $(\mathcal{W}^d, \mathcal{B}(\mathcal{W}^d), \mathcal{N}_t, P_{s,x})$, on which the family of martingales $M_t^i := x_t^i - x_0^i - \int_0^t \beta^i(s, x_s)\, ds$ are given by a stochastic integral w.r.t a Wiener process $\tilde{w}_t$ defined on this extended probability space.

First we have the technical lemma on measurability.

Lemma 4.2.23. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space. Suppose $a : [0, T] \times \Omega \to S_+^d$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}$ measurable and $\mathcal{F}_t$−adapted. Then $\sqrt{a} \in S_+^d$ is also measurable and $\mathcal{F}_t$ adapted.*

Proof. First consider the case a is in diagonal form, the diagonal has non negative entries and without loss of generality (the other cases are treated similarly) suppose that it is given by $a = diag(\lambda_1, \ldots \lambda_l, 0, \ldots 0)$ with $\lambda_i > 0$ for $i = 1, 2, \ldots, l$ with $l \leqslant d$, a simple computation (on each entry) yields

$$\sqrt{a_t} = c \int_0^t \frac{I - e^{-a_t \cdot u}}{u^{3/2}} du,$$

Fubini theorem implies the required measurability properties in this case.
In the case $a(s,\omega)$ is not a diagonal matrix, by the properties of positive semi-definite matrices, for each $(t,\omega) \in [0,\infty) \times \Omega$ there is a decomposition $a_t = Q_t \cdot D_t \cdot Q_t'$ with Q_t is an orthogonal matrix and D is a diagonal matrix

$$D = \begin{pmatrix} \lambda_1 & 0 & . & . & . & \dots & \mathbf{0} \\ 0 & \lambda_2 & 0 & . & . & \dots & \mathbf{0} \\ 0 & 0 & \lambda_3 & . & . & \dots & \mathbf{0} \\ 0 & 0 & 0 & \lambda_4 & . & \dots & \mathbf{0} \\ 0 & 0 & 0 & 0 & \lambda_5 & \dots & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & \dots & \mathbf{0} \\ 0 & 0 & \cdot & \cdot & \cdot & \dots & \mathbf{0}_{d-l} \end{pmatrix}.$$

Where $\mathbf{0}_{d-l} \in \mathbb{R}^{(d-1)\times(d-l)}$ with all its entries equal to 0. In this case we can obtain a similar decomposition of $\sqrt{a_t}$ given by $\sqrt{a_t} = Q_t \cdot \sqrt{D_t} \cdot Q_t'$ satisfies the condition. Thus,

$$Q_t\sqrt{D_t}Q_t' = c\int_0^\infty \frac{I - Q_t \cdot e^{-D_t \cdot u}Q_t'}{u^{3/2}}du = c\int_0^\infty \frac{I - e^{-a_t \cdot u}}{u^{3/2}}du,$$

by properties of the exponential of the matrix and the fact that Q_t is orthogonal.

□

Lemma 4.2.24. *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space whose filtration satisfies the usual conditions. Suppose $a_s(\omega)$ is a symmetric and positive definite matrix $\mathbb{P}-a.s.$ for every $s \geqslant 0$. Let $\left(\xi_s^i\right)_{s\geqslant 0}$ be a $d-$dimensional local martingale such that $\xi_t^i \cdot \xi_t^j - \int_0^t a_s^{ij} ds$ is a local martingale for $i,j = 1,\dots,n$, then there is a $d-$dimensional $\mathcal{F}_t-$Wiener process $\tilde{w}_s$ and an $\mathcal{F}_t-$adapted and measurable process f_s such that*

$$\xi_s = \xi_0 + \int_0^t f_s d\tilde{w}_s \ \mathbb{P}-a.s.$$

Proof. As a_s is symmetric and positive definite, a_s is diagonalisable $(s,\omega)-a.e.$ The eigenvalues are given by the set $\{\lambda \in \mathbb{R} : det(a(s,\omega) - \lambda I) = 0\} = \{\lambda_1(s,\omega), \lambda_2(s,\omega), \dots, \lambda_d(s,\omega)\}$ to see that the eigenvalues are measurable, one can consider the set-valued mapping $\Lambda : \mathbb{R}^+ \times \Omega \longmapsto \mathbb{R}^d$, a theorem on the measurability of set-valued mappings[6] states that it is enough to see that the graph of Λ is jointly measurable,

$$gph(\Lambda) = \{(s,\omega,x) : det(a(s,\omega) - x\mathbb{I}_d) = 0\} \in \mathcal{B}\left(\mathbb{R}^+\right) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{R}),$$

as the determinant function $det(\cdot)$ is continuous and $a(s,\omega)$ is $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^+)-$measurable our claim follows. The same argument shows that the eigenvalues $\lambda_1(t,\cdot),\dots,\lambda_d(t,\cdot)$ are also $\mathcal{F}_t-$adapted. Similarly, the set $\left\{v \in \mathbb{R}^d : a_s(\omega)v = \lambda_s \cdot v\right\}$ is also $\mathcal{F}_t-$adapted and $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^+)$-measurable. Hence, the orthogonal matrices Q_t and Q_t' in the decomposition of the symmetric matrix $a_s = Q_s D_s Q_s'$ are measurable and $\mathcal{F}_t-$adapted, where $D_s = diag\{\lambda_1(s),\dots,\lambda_d(s)\}$. The matrix $\sqrt{a_s}$ is invertible and its inverse is also measurable hence $\sqrt{D_t}$ is well defined and it is invertible. Then $\left(\sqrt{a_s}\right)^{-1} = Q_s\sqrt{D_s}^{-1}Q_s'$ has the same measurability properties, as well.

[6]See appendix C

Define

$$f_s = \sqrt{a_s}^{-1} \cdot \mathbb{I}_{(det(a_s)>0)} + I_d \cdot \mathbb{I}_{(det(a_s)=0)}, \tag{4.2.32}$$

Where I_d is the identity matrix in $\mathbb{R}^d$. Then $\sqrt{a_s} \cdot f_s = I_d \mathbb{I}_{det(a_s)\neq 0}$ $l \times \mathbb{P} - a.e.$ by assumption on a_s here $l \times \mathbb{P}$ denotes the product measure, on $\mathbb{R}^+ \times \Omega$, of the Lebesgue measure l (restricted to the Borel sets) and the probability measure $\mathbb{P}$.
In this case $\sqrt{a_s} f_s = I_d$ $l \times \mathbb{P} - a.e.$, therefore, $\sum_{k=1}^d \left(\sqrt{a_s} f_s\right)^{ik} \left(\sqrt{a_s} f_s\right)^{jk} = \delta_{ij}$ and

$$\sum_{k,l=1}^d \int_0^t f_s^{ik} f_s^{jl} a_s^{kl} ds = \sum_{k,l=1}^d \sum_{p=1}^d \int_0^t f_s^{ik} f_s^{jl} \left(\sqrt{a_s}\right)^{kp} \left(\sqrt{a_s}\right)^{pl} ds,$$

as the matrices are symmetric, rearranging

$$\sum_{k,l=1}^d \int_0^t f_s^{ik} f_s^{jl} a_s^{kl} ds = \sum_{p=1}^d \int_0^t \left(f_s \sqrt{a_s}\right)^{ip} \left(f_s \sqrt{a_s}\right)^{pj} ds = t\delta^{ij} < \infty, \tag{4.2.33}$$

$$\int_0^t tr\left(\sqrt{a_s} f_s f_s' \sqrt{a_s}'\right) ds = \sum_{i=1}^d \int_0^t \left(\sqrt{a_s} f_s f_s' \sqrt{a_s}'\right)^{ii} ds = t < \infty.$$

As (4.2.33) implies that $f_s \in \mathscr{P}(\langle M \rangle)$[7], thus, the process $\int_0^t f_s dx_s$ is a $\left(\mathcal{F}_t, \mathbb{P}\right)$ local martingale and

$$\left\langle \left(\int_0^\cdot f_s dx_s\right)^{(i)}, \left(\int_0^\cdot f_s dx_s\right)^{(j)} \right\rangle_t = \sum_{k=1}^d \sum_{l=1}^d \int_0^t f_s^{ik} f_s^{jl} d\langle x_k, x_l \rangle_s = \delta^{ij} t.$$

By Lévy's theorem then $\tilde{w}_t = \int_0^t f_s dx_s$ is an $d-$dimensional Wiener process. Finally, by properties of the stochastic integrals

$$\int_0^t \sqrt{a_s} d\tilde{w}_s = \int_0^t \sqrt{a_s} f_s dx_s = \int_0^t \sqrt{a_s} \left(\sqrt{a_s}^{-1} \cdot \mathbb{I}_{(det(a_s)>0)} + I_n \cdot \mathbb{I}_{det(a_s)=0}\right) dx_s = x_t - x_0, \tag{4.2.34}$$

by the fact that, under our assumptions, $\mathbb{I}_{(det(a_s)>0)} = 1,\ l \times \mathbb{P}-\ a.e.$ This proves the lemma. □

One can see from the proof how crucial was the assumption $\mathbb{I}_{(det(a_s)>0)} = 1$ $\mathbb{P} - a.s.$ for all $s \geqslant 0$ in order to apply Lévy's theorem.
It is possible to prove an analogous result in the case of a *degenerate* diffusion, i.e. $\det(a_s(\omega)) = 0$, has positive $l \times \mathbb{P}-$measure, however an extension of the probability space is required if one aims to obtain a process whose quadratic variation is equal to $\delta^{ij} t$.
For this purpose we introduce the concept of an extension of a filtered probability space

Definition 4.2.25. We say that the complete filtered probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \left\{\widetilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \widetilde{\mathbb{P}}\right)$ is an extension of the filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ if $\left(\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\geqslant 0}, \mathbb{P}'\right)$ is another filtered probability space and

- $\widetilde{\Omega} = \Omega \times \Omega'$,
- $\widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$,

[7]The class of measurable and $\mathcal{F}_t-$adapted processes $\left(X_t\right)_{t\geqslant 0}$ satisfying $\mathbb{P}\left(\omega : \int_0^T X_t^2 d\langle M \rangle_t (\omega) < \infty\right) = 1$ for all $T > 0$

- $\widetilde{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'_t$, and
- $\widetilde{\mathbb{P}} = \mathbb{P} \times \mathbb{P}'$.

On $\widetilde{\Omega}$ define the projections $p : \widetilde{\Omega} \to \Omega$ and $p' : \widetilde{\Omega} \to \Omega'$ by $p(\tilde{\omega}) = \omega$ and $p'(\tilde{\omega}) = \omega'$, for $\tilde{\omega} = (\omega, \omega') \in \widetilde{\Omega}$.

There is a more general concept of an extension of a filtered probability space, but we shall not use it throughout the proof. See Definition II.7.1 in [25].

On this extended stochastic basis, it is possible to define 'extensions' of the processes previously defined on either $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ or $\left(\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\geqslant 0}, \mathbb{P}'\right)$. Indeed, define $\tilde{x} : \mathbb{R}_+ \times \widetilde{\Omega} \to \mathbb{R}^d$ and $\tilde{B} : \mathbb{R}_+ \times \widetilde{\Omega} \to \mathbb{R}^d$ by

$$\tilde{x}(t, \tilde{\omega}) := x(t, p\tilde{\omega}) \text{ and } \tilde{B}(t, \tilde{\omega}) := B(t, p\tilde{\omega}). \tag{4.2.35}$$

This simply means $\tilde{x}(t, \omega, \omega') = x(t, \omega)$ for all $t \geqslant 0$ and all $\omega' \in \Omega'$, and $\tilde{B}(t, (\omega, \omega')) = B(t, \omega')$ for all $t \geqslant 0$ and all $\omega \in \Omega$. Similarly, we can extend the S^d_+−valued process $a(s, \omega)$ to $\tilde{a}(s, \tilde{\omega})$, by defining $\tilde{a}(s, \tilde{\omega}) := a(s, p\tilde{\omega})$. The projections p and p' are measurable, for any $B \in \mathcal{F}_t$ we have $p^{-1}(B) \in \widetilde{\mathcal{F}}_t$ as $\{\tilde{\omega} : p(\tilde{\omega}) \in B\} = B \times \Omega'$.

Furthermore, we have the following identity between the conditional expectations of random variables defined on this extended probability space. Let $\tilde{Z}(\tilde{\omega}) \in L^1\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ be a r.v. on $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ and such that $\tilde{Z}$ is an extension of a r.v. $Z(\omega)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ then

$$\tilde{\mathbf{E}}\left[\tilde{Z}\middle|\widetilde{\mathcal{F}}_t\right](\tilde{\omega}) = \mathbf{E}\left[Z\,|\mathcal{F}_t\right](p\tilde{\omega}). \tag{4.2.36}$$

The proof of this fact is a consequence of Fubini's theorem, indeed, let $A \times B \in \mathcal{F}_t \otimes \mathcal{F}'_t$ then

$$\int_{A\times B} \tilde{\mathbf{E}}\left[\tilde{Z}\middle|\widetilde{\mathcal{F}}_t\right](\tilde{\omega})\, d\widetilde{\mathbb{P}}(\tilde{\omega}) = \int_{\Omega'}\int_{\Omega} \mathbb{I}_A(\omega)\, \mathbb{I}_B(\omega')\, \tilde{Z}(\omega, \omega')\, d\mathbb{P}(\omega)\, d\mathbb{P}'(\omega') =$$

$$= \int_{\Omega'} \mathbb{I}_B(\omega')\, d\mathbb{P}'(\omega') \int_{\Omega} \mathbb{I}_A(\omega)\, Z(\omega)\, d\mathbb{P}(\omega) = \int_{\Omega'} \mathbb{I}_B(\omega')\, d\mathbb{P}'(\omega') \int_{\Omega} \mathbb{I}_A(\omega)\, \mathbf{E}\left[Z\middle|\mathcal{F}_t\right] d\mathbb{P}(\omega) = \tag{4.2.37}$$

$$= \int_{\Omega\times\Omega'} \mathbb{I}_A(\omega)\, \mathbb{I}_B(\omega')\, \mathbf{E}\left[Z\,|\mathcal{F}_t\right](\omega)\, d\mathbb{P}(\omega)\, d\mathbb{P}'(\omega') = \int_{A\times B} \mathbf{E}\left[Z\,|\mathcal{F}_t\right](p\tilde{\omega})\, d\widetilde{\mathbb{P}}(\tilde{\omega}). \tag{4.2.38}$$

As the r.v. $\mathbf{E}\left[Z\,|\mathcal{F}_t\right](p\tilde{\omega})$ is $\widetilde{\mathcal{F}}_t$−measurable (a composition of an $\mathcal{F}_t$−measurable function and a projection), then (4.2.36) holds.

The identity in (4.2.36) also implies that given a martingale $\{X_t\}_{t\geqslant 0}$ with respect to $(\mathcal{F}_t, \mathbb{P})$ its extension given by $\left\{\tilde{X}_t\right\}_{t\geqslant 0}$ is also a $\left(\widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$ martingale. In particular, if $M \in \mathcal{M}^2(\mathcal{F}_t, \mathbb{P})$ (a square integrable martingale) then the extended process $\tilde{M} \in \mathcal{M}^2\left(\widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$ and hence one can also check that for $M, N \in \mathcal{M}^2\left(\widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$ it follows that $\langle M, N\rangle_t(p\tilde{\omega}) = \left\langle \tilde{M}, \tilde{N}\right\rangle_t(\tilde{\omega})$ therefore, the spaces $\mathcal{M}^2$, $\mathcal{M}^{2,c}$ and $\mathcal{M}^2_{loc}$ are imbedded into the spaces $\mathcal{M}^2\left(\widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$, $\mathcal{M}^{2,c}\left(\widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$ and $\mathcal{M}^2_{loc}\left(\widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}}\right)$, respectively. Moreover, it is easy to verify that the identity (4.2.35) also implies that the extension $\{\tilde{B}_t\}_{t\geqslant 0}$ is an $\left(\tilde{\mathcal{F}}_t, \tilde{\mathbb{P}}\right)-$Wiener process.

Bearing in mind these facts we obtain a generalisation of Lemma 4.2.24.

Proposition 4.2.26. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space whose filtration satisfies the usual conditions. Let $a : \Omega \times \mathbb{R}^+ \to S_+^d$, be a positive semi-definite matrix, and let $\left(\xi_s^i\right)_{s\geqslant 0}$ be a $d-$dimensional admissible local martingale with $a^{ij}(s,\omega)$ satisfying (4.2.12). Denote by $\sqrt{a_s}$ the square root of the matrix a_s, in particular, $\sqrt{a_s} \in \mathcal{A}^{d,d}$. Then there is an extension of the probability space, in the sense of Definition 4.2.25, denoted by $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$, an extension of the processes into this extended filtered probability space and a $\tilde{\mathcal{F}}_t-$ Wiener process such that*

$$\tilde{\xi}_t(\tilde{\omega}) - \tilde{\xi}_0(\tilde{\omega}) = \int_0^t \sqrt{\tilde{a}_s(\tilde{\omega})} d\tilde{w}_s(\tilde{\omega}) \quad \tilde{\mathbb{P}} - a.s. \text{ for all } t \geqslant 0, \tag{4.2.39}$$

Proof. Let $(\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\geqslant 0}, \mathbb{P}')$ be a filtered probability space, whose filtration satisfies the usual conditions and such that there is an $\mathcal{F}'_t-$Wiener process $\{B_t\}_{t\geqslant 0}$ on $(\Omega', \mathcal{F}', \mathbb{P}')$. Define the extension of the probability space, $\tilde{\Omega} = \Omega \times \Omega'$, $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$, $\tilde{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'_t$ and $\tilde{\mathbb{P}} = \mathbb{P} \times \mathbb{P}'$. On $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}, \tilde{\mathbb{P}}\right)$ we can define measurable 'extensions' of the stochastic processes $\tilde{\xi}_s(\tilde{\omega}) := \xi_s(\omega)$, $\tilde{a}_s(\tilde{\omega}) = a_s(\omega)$ i.e. all the processes previously defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$.
Consider the extensions of the processes $a(s,\omega)$ and $\xi_s(\omega)$ as explained after Definition 4.2.25. On $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$ the following process

$$f_s^\varepsilon = \left(\sqrt{\tilde{a}_s} + \varepsilon^2 \cdot I\right)\left(\tilde{a}_s + \varepsilon I\right)^{-1}, \text{ and } f_s = \lim_{\varepsilon\to 0} f_s^\varepsilon = \lim_{\varepsilon \to 0}\left(\sqrt{\tilde{a}_s} + \varepsilon^2 \cdot I\right)\left(\tilde{a}_s + \varepsilon I\right)^{-1}.$$

We claim that $f_s\sqrt{\tilde{a}_s}$ is the orthogonal projection over the (random) subspace $Range\left(\sqrt{\tilde{a}_s}\right)$. As we can decompose $\eta \in \mathbb{R}^d$ as a sum of $\eta_0 \in Range\left(\sqrt{\tilde{a}_s}\right)$ and $\eta_\perp \in \left[Range\left(\sqrt{\tilde{a}_s}\right)\right]^\perp = Ker\left(\sqrt{\tilde{a}_s}'\right) = Ker\left(\sqrt{\tilde{a}_s}\right)$ we consider each of these cases. Suppose $\eta \in \left[Range\left(\sqrt{\tilde{a}_s}\right)\right]^\perp$ then $f_s\tilde{a}_s\eta = 0$. If $\eta \in Range\left(\sqrt{\tilde{a}_s}\right)$, then there is a $z \in \mathbb{R}^d$ with $\eta = \sqrt{\tilde{a}_s}z$ and $f_s\sqrt{\tilde{a}_s}\eta = f_s\tilde{a}_s z$.

Define $\pi_\varepsilon = f_s^\varepsilon\sqrt{\tilde{a}_s}$ then $\pi_\varepsilon\eta = \pi_\varepsilon\sqrt{\tilde{a}_s}z = f_s^\varepsilon\tilde{a}_s z = f_s^\varepsilon(\tilde{a}_s + \varepsilon I)z - \varepsilon f_s^\varepsilon z = \left(\sqrt{\tilde{a}_s} + \varepsilon^2 I\right)z - \varepsilon f_s^\varepsilon z = \eta - \varepsilon\left((\varepsilon I - f_s^\varepsilon)z\right) \to \eta$. Thus $\lim_{\varepsilon\to 0}\pi_\varepsilon = \pi_a$ where π_a is the orthogonal projection onto the $Range\left(\sqrt{a_s}\right)$.
Define $g_s = I - f_s\sqrt{a_s} = \pi_{[Range(\sqrt{a_s})]^\perp}$.

$$f_s\sqrt{a_s} = \pi_{Range(\sqrt{a_s})}. \tag{4.2.40}$$

As was pointed out, the measurability properties of a_s are preserved by $\sqrt{a_s}$ and the same holds for the extension $\sqrt{\tilde{a}_s}$. Following the proof in Lemma 4.2.24 if we define

$$w'_t(\tilde{\omega}) := \int_0^t \tilde{f}_s(\tilde{\omega})\, d\tilde{\xi}_s,$$

it follows that the process w'_t is a local martingale if the integrand is such that $\int_0^t tr\left(\tilde{f}_s\tilde{a}_s\tilde{f}'_s(\tilde{\omega})\right)ds < \infty$ $\tilde{\mathbb{P}}-a.s.$, here $\tilde{f}'_s$ is the transpose of the matrix valued process f_s. But $\tilde{f}_s\tilde{a}_s\tilde{f}'_s = \left(\tilde{f}_s\sqrt{\tilde{a}_s}\right)\left(\tilde{f}_s\sqrt{\tilde{a}_s}\right)' = \pi_{\tilde{a}} \cdot \pi_{\tilde{a}} = \pi_{\tilde{a}}$ by properties of orthogonal projections.

$$\int_0^t \sum_{k=1}^d \sum_{l=1}^d \tilde{f}_s^{ik}\tilde{a}_s^{kl}\tilde{f}_s^{jl} ds \leqslant \int_0^t (\pi_{\tilde{a}})^{ij} ds < \infty,$$

as for any orthogonal projections we have $\|\pi_{\tilde{a}_s}\| \leqslant 1$.

Then w'_t is a local martingale (each of its entries is a local martingale). And

$$\left\langle (w')^i, (w')^j \right\rangle_t = \int_0^t (\pi_{a_s})^{ij} \, ds. \tag{4.2.41}$$

Unlike the proof of Lemma 4.2.24 w'_t is not a Wiener process on $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$ as *Range* $\left(\sqrt{\tilde{a}_s}\right)$ is a proper subspace, hence, (4.2.41) will never be the identity matrix.
Define on $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$ the following process

$$\tilde{w}_t(\tilde{\omega}) = w'_t(\tilde{\omega}) + \int_0^t \left(I - \tilde{f}_s\sqrt{\tilde{a}_s}\right) d\tilde{B}_s(\tilde{\omega}) \tag{4.2.42}$$

$$= \int_0^t f_s(\tilde{\omega})\, d\tilde{\xi}_s + \int_0^t \left(I - \tilde{f}_s\sqrt{\tilde{a}_s}(\omega)\right) d\tilde{B}_s(\tilde{\omega}).$$

One can check that for fixed (s, ω) and $\int_0^t \left(\left(I - \tilde{f}_s\sqrt{\tilde{a}_s}\right)\left(I - \tilde{f}_a\sqrt{\tilde{a}_s}\right)'\right)^{ij} ds < \infty$, for all $i, j \leqslant d$ hence the second integral is a local martingale. The process $\tilde{w}$ is a local martingale as it is the sum of two well-defined stochastic integrals on $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}}\right)$. Moreover, it is square integrable

$$\tilde{\mathbf{E}}\left\| \int_0^t f_s(\tilde{\omega})\, d\tilde{\xi}_s \right\|^2 \leqslant 2 \cdot \mathbf{E}\left\| w'_t(\tilde{\omega}) \right\|^2 + 2 \cdot \mathbf{E}'\left\| \int_0^t (I - f_s\sqrt{a_s})\,(p(\tilde{\omega}))\, d\tilde{B}_s\left(p'\tilde{\omega}\right) \right\|^2.$$

We write $\tilde{B}_s(p'\tilde{\omega})$ as $B_s(\omega')$, i.e. the process $\tilde{B}$ is equal to $B_s(\omega')$ for a given $\tilde{\omega}= (\omega, \omega')$. Something similar can be said about $\tilde{w}_t$.
It follows from previous computations (see 4.2.41) that the first stochastic integral in (4.2.42) is in fact a martingale and

$$\tilde{\mathbf{E}}\left\| \int_0^t f_s(\tilde{\omega})\, d\tilde{\xi}_s \right\|^2 = \tilde{\mathbf{E}}\Big\langle \int_0^{\cdot} \tilde{f}_s d\tilde{\xi}_s \Big\rangle_t = \int_0^t tr\,(\pi_{a_s})\, ds \leqslant td < \infty,$$

as orthogonal are such that $\|\pi_{\tilde{a}_s}\| \leqslant 1$ (here the bound is obtained when regarding π as an operator). Similarly, (as $I - \pi_{\tilde{a}_s}$ is also an orthogonal projection)

$$\mathbf{E}'\| \int_0^t \left(I - \tilde{f}_s\sqrt{\tilde{a}_s}\right) d\tilde{B}_s\|^2 < \infty.$$

Similar computations yield

$$\left\langle \tilde{w}^i, \tilde{w}^j \right\rangle_t(\tilde{\omega}) = \int_0^t \left(\pi_{\tilde{a}_s}\pi^{\star}_{\tilde{a}_s}\right)^{ij} ds + \int_0^t \left((I - \pi_{\tilde{a}_s})(I - \pi_{\tilde{a}_s})^{\star}\right)^{ij} ds = \delta^{ij} t. \tag{4.2.43}$$

By Lévy's theorem, $\tilde{w}_s$ defined in (4.2.42) is an $\tilde{\mathcal{F}}_t$-Wiener process on the filtered probability space $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}}\right)$.
We aim to represent $\tilde{x}$, as an integral w.r.t. to $\tilde{w}_s$ *on* $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}}\right)$.
We can readily check that $\sqrt{\tilde{a}_s} f_s = \pi_{range\left(\sqrt{\tilde{a}_s}\right)}$ (also denoted by $\pi_{\tilde{a}_s}$). The proof of this fact follows the steps of the previous verification for the process $f_s\sqrt{a_s}$ (reduce to check for $\eta \in$ *Range* $\left(\sqrt{\tilde{a}_s}\right)$ and its orthogonal complement, so it is omitted).

$$\int_0^t \sqrt{\tilde{a}_s} d\tilde{w}_s = \int_0^t \sqrt{\tilde{a}_s}\tilde{f}_s d\tilde{\xi}_s(\tilde{\omega}) + \int_0^t \sqrt{\tilde{a}_s}\,(I - \pi_{\tilde{a}_s})\, d\tilde{B}_s(\tilde{\omega}), \tag{4.2.44}$$

$$\int_0^t \sqrt{\tilde{a}_s} d\tilde{w}_s = \int_0^t \sqrt{a_s} f_s(\omega)\, d\xi_s(\omega)\,. \tag{4.2.45}$$

As $\sqrt{a_s}\pi_{a_s} = \sqrt{a_s}$ by definition of π_{a_s}, and $\sqrt{\tilde{a}_s}\left(I - \pi_{\tilde{a}_s}\right) = 0$. It follows that

$$\int_0^t \sqrt{\tilde{a}_s}\tilde{f}_s d\tilde{\xi}_s(\tilde{\omega}) = \int_0^t \pi_{a_s} d\xi_s = \xi_s - \xi_0. \tag{4.2.46}$$

The claim will be proved if we show that last equalities in (4.2.46) holds.
Notice that

$$\left\langle \left(\pi_{a_s}\xi_\cdot\right)^i, \left(\pi_{a_s}\xi_\cdot\right)^j \right\rangle_t = \left\langle \sum_{k=1}^d \pi_{a_s}^{ik}\xi^k, \sum_{l=i}^d \pi_{a_s}^{jl}\xi^l \right\rangle_t = \tag{4.2.47}$$

$$\sum_{k=1}^d \sum_{l=1}^d \pi_{a_s}^{ik}\pi_{a_s}^{jl}\left\langle \xi^k, \xi^l \right\rangle_s = \int_0^t \sum_{k=1}^d \sum_{l=1}^d \pi_{a_s}^{ik}\pi_{a_s}^{jl} a_s^{kl} ds = \int_0^t \left(\pi_{a_s} a_s \pi_{a_s}\right)^{ij} ds = \int_0^t a_s^{ij} ds,$$

since $\pi_{a_s} a_s \pi_{a_s} = \pi_{a_s}\sqrt{a_s}\sqrt{a_s}\pi_{a_s} = \sqrt{a_s}\sqrt{a_s}\pi_{a_s} = a_s$ again, by an analogous argument in the verification of the identity $\tilde{f}_s\sqrt{\tilde{a}_s}$ and the definition of the orthogonal projection. Thus,

$$\left\langle \left(\pi_{a_s}\xi\right)^i, \left(\pi_{a_s}\xi\right)^j \right\rangle_t = \int_0^t a_s^{ij} ds. \tag{4.2.48}$$

Similarly one can check the following,

$$\left\langle \xi_\cdot^i - \xi_0^i - \left(\int_0^\cdot \pi_{a_s} d\xi_s\right)^i, \xi_\cdot^j - \xi_0^j - \left(\int_0^\cdot \pi_{a_s} d\xi_s\right)^j \right\rangle_t = \left\langle \left(\int_0^\cdot \left(I - \pi_{a_s}\right) d\xi_s\right)^i, \left(\int_0^\cdot \left(I - \pi_{a_s}\right)\right)^j d\xi_s \right\rangle_t, \tag{4.2.49}$$

by properties of the quadratic variation we have

$$\left\langle \left(\int_0^\cdot \left(I - \pi_{a_s}\right) d\xi_s\right)^i, \left(\int_0^\cdot \left(I - \pi_{a_s}\right)\right)^j d\xi_s \right\rangle_t = \int_0^t \left(\pi_{Range(a_s)^\perp} a_s\right)^{ij} ds = 0.$$

From this it follows that $\xi_t - \xi_0 = \int_0^t \pi_{a_s} d\xi_s = \int_0^t \tilde{f}_s d\tilde{w}_s$.
And the theorem is proved.

□

The next proposition settles up many facts concerning measurability, and states a decomposition of the matrix a_s. For a proof of this lemma please refer to [52].

Proposition 4.2.27. *Let a be an element of S_+^d, if π is the orthogonal projection into the subspace $Range(a) \subset \mathbb{R}^d$ then there is a matrix $\tilde{a} \in S_+^d$ such that $\tilde{a}a = \pi = a\tilde{a}$ moreover*

$$\pi = \lim_{\epsilon \to 0}\left(a + \epsilon I\right)^{-1} a \text{ and } \tilde{a} = \lim_{\epsilon \to 0}\left(a + \epsilon I\right)^{-1}\pi_\sigma.$$

Suppose $\sigma \in \mathbb{R}^{d\times m}$ and such that $\sigma \cdot \sigma' = a$. Let π_σ denote the orthogonal projection into $Range(\sigma') \subset \mathbb{R}^d$ then $Range(a) = Range(\sigma)$ and $\sigma' \cdot \tilde{a}\sigma = \pi_\sigma$.

Remark 4.2.28. Again, $\tilde{a}$ resembles the inverse of a. The approximation can be done in different ways. For instance, one can see that $\tilde{a} = \lim_{\epsilon\to 0}\left(a + \epsilon^2 I\right)\left(a^2 + \epsilon I\right)^{-1} = \lim_{\epsilon\to 0}\tilde{a}_s^\epsilon$. By the above-mentioned decomposition of $\mathbb{R}^d$ in orthogonal subspaces one can readily check that $\lim_{\epsilon\to 0}\tilde{a}_s^\epsilon a_a = \pi = \lim_{\epsilon\to 0} a_s\tilde{a}_s^\epsilon$.

Using Proposition 4.2.27 and a similar argument as in Theorem 4.2.24 we prove the following theorem

Theorem 4.2.29. *Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ be a filtered probability space whose filtration satisfies the usual conditions, let $\{\xi_t\}_{t\geqslant 0}$ be a stochastic process as in Proposition 4.2.26. Let $a : [0,\infty) \times \mathbb{R}^d \to S_+^d$ and $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be bounded measurable functions, suppose*

$$\xi_t - \xi_0 - \int_0^t b(s, \xi_s)\, ds \in \mathcal{M}_{loc}(\mathcal{F}_t, \mathbb{P}), \tag{4.2.50}$$

such that

$$\left(\xi_t - \xi_0 - \int_0^t b(s,\xi_s)\right)^i \cdot \left(\xi_t - \xi_0 - \int_0^t b(s,\xi_s)\right)^j - \int_0^t a^{ij}(s,\xi_s)\, ds,$$

is a local martingale for each $i, j = 1, \ldots, d$, in other words, $m_t = \xi_t - \xi_0 - \int_0^t b(s,\xi_s)\, ds$ is an admissible local martingale. Let $\sigma : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ with $a = \sigma \cdot \sigma'$ then there is an extension of the filtered probability space $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$ and a $\tilde{\mathcal{F}}_t-$ Wiener process $\tilde{w}_t(\tilde{\omega})$ such that

$$\tilde{\xi}_t = \tilde{\xi}_0 + \int_0^t b\left(s, \tilde{\xi}_s\right) ds + \int_0^t \sigma\left(s, \tilde{\xi}_s\right) d\tilde{w}_s \ \ \mathbb{P} - a.s.$$

i.e. $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}, \{\tilde{\xi}_t\}_{t\geqslant 0}, \{\tilde{w}_t\}_{t\geqslant 0}\right)$ is a weak solution and the law of ξ is given by $\tilde{\mathbb{P}} \circ \varphi_{\tilde{\xi}}^{-1}(\cdot)$.

These results allow to prove the equivalence of the solution a stochastic differential equation and the solution to the martingale problem.

Theorem 4.2.30. *Let $\left(\mathcal{W}^d, \mathcal{N}, \{\mathcal{N}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ be the canonical space and $\mathcal{N} = \mathcal{B}\left(\mathcal{W}^d\right)$. Suppose $\mathcal{N}_t = \sigma(\alpha_s,\, s \leqslant t)$, and P_0 is a solution to the martingale problem $(0, \mu, \mathcal{L}_t)$, then there is an extension of the previous 'stochastic basis' $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$, an $\tilde{\mathcal{F}}_t-$ Wiener process, $\tilde{w}_t$ and an $\tilde{\mathcal{F}}_t-$adapted, measurable stochastic process $\tilde{x}_t$ on the extended probability space such that*

$$dx_t = b(t, x_t)\, dt + \sigma(t, x_t)\, dw_t, \tag{4.2.51}$$

$$x(0) \sim \mu.$$

holds for $(\tilde{w}_t, \tilde{x}_t)_{t\geqslant 0}$.
Conversely, suppose μ is a Borel measure on $\mathcal{B}\left(\mathbb{R}^d\right)$ with $\mathbb{P}(X_0 \in \cdot) = \mu(\cdot)$, and $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}, x, w\right)$ is a weak solution of (4.2.51) with $X(0) \sim \mu$ i.e. the initial condition $X(0)$ has law μ. Then the law of the process X, $\mathbb{Q} = \mathbb{P} \circ X^{-1}$ is a solution to the martingale problem $(0, \mu, \mathcal{L}_t)$.

Proof. Suppose P_0 is the solution of the martingale problem and denote by x_t the canonical process on $\left(\mathcal{W}^d, \mathcal{N}, \{\mathcal{N}_t\}_{t\geqslant 0}, P_0\right)$, define

$$M_t^y = \left\langle y, x_t - x_0 - \int_0^t b(u, x_u)\, du \right\rangle,$$

by Proposition 4.2.20, M_t^y is a $(\mathcal{N}_t, P_0)$ local martingale and $\langle M^y \rangle_t = \int_0^t \langle y, a(s, x_s)\, y \rangle ds$ for all $t \geqslant 0$. By the last proposition, there is an extension $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$ (possibly trivial, if

a_s is non-singular $\lambda \times \mathbb{P} - a.e.$) and a $\tilde{\mathcal{F}}_t-$ d-dimensional Wiener process $\tilde{w}_t$ such that

$$\tilde{M}_t(\tilde{\omega}) := \int_0^t \sigma(s, \tilde{x}_s(\tilde{\omega}))\, d\tilde{w}_s(\tilde{\omega}).$$

In other words, there is a filtered probability space $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \mathbb{P}\right)$ and stochastic processes $(\tilde{x}, \tilde{w})$ (adapted and measurable) such that

$$\tilde{x}_t - \tilde{x}_0 = \int_0^t b(u, \tilde{x}_u)\, du + \int_0^t \sigma(u, \tilde{x}_u)\, d\tilde{w}_u,\ \tilde{\mathbb{P}} - a.s.$$

Now suppose that there is an stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ and stochastic processes (X_t, W_t) such that (4.2.51) holds and W_t is an $\mathcal{F}_t-$Wiener process. If $f \in C_0^\infty(\mathbb{R}^d;\mathbb{R})$, let $0 \leqslant s < t$ Itô's formula yields

$$f(X_t) = f(X_s) = \int_s^t \mathcal{L}_u f(X_u)\, du + \int_s^t \nabla f(X_u)' \sigma(u, X_u)\, dW_u. \tag{4.2.52}$$

Thus, the process

$$C_t^{s,f} = f(X_t) - f(X_s) - \int_s^t \mathcal{L}_u f(X_u)\, du,$$

is an $(\mathcal{F}_t, \mathbb{P})-$martingale. Denoting by $\{x_s\}_{s\geqslant 0}$ the canonical process on $\left(\mathcal{W}^d, \mathcal{N}, \mathbb{P}\circ(\varphi_X)\right)$, we claim that the martingale property holds as well for the process $C_t^{s,f}$ defined on the canonical space, that is, we claim that $f(x_t(\omega')) - f(x_s(\omega')) - \int_s^t \mathcal{L}_u f(x_u(\omega'))\, du$ is an $\left(\mathcal{N}_t, \mathbb{P}\circ(\varphi_X)^{-1}\right)$ -martingale. Indeed, if $A \in \mathcal{N}_s$ without loss of generality we can assume $A = x_r^{-1}(B)$ with $B \in \mathcal{B}(\mathbb{R}^d)$ and $r \leqslant s$ then we aim to compute

$$\int_A f\left(x_t(\omega')\right) - f(x_s(\omega)) - \int_s^t \mathcal{L}_u f\left(x_u(\omega')\right) du d\mathbb{P}\circ\varphi_X\left(\omega'\right) = \tag{4.2.53}$$

$$= \int_{\varphi^{-1}(A)} f(x_t(\varphi_X(\omega))) - f(x_s(\varphi_X(\omega))) - \int_s^t \mathcal{L}_u f(x_u(\varphi_X(\omega)))\, du d\mathbb{P}(\omega), \tag{4.2.54}$$

notice that the set $\varphi_X^{-1}(A)$ is $\mathcal{F}_t-$measurable, in fact

$$\varphi_X^{-1}(A) = (x_r \circ \varphi_X)^{-1}(B) = \{\omega : X_r(\omega) \in B\} \in \mathcal{F}_t.$$

On the other hand, $f(x_u(\varphi_X(\omega))) = f(X_u(\omega))$, thus (4.2.54) is equal to

$$\int_{\varphi^{-1}} (A) f(X_t(\omega)) - f(X_s(\omega)) - \int_s^t \mathcal{L}_u f(X_u(\omega))\, du d\mathbb{P}(\omega) = 0.$$

As $C_t^{s,f}$ is a $(\mathcal{F}_t, \mathbb{P})-$martingale. Moreover, $\mathbb{P}(X_0 \in B) = \mu(B) = \left(\mathbb{P}\circ\varphi_X^{-1}\right)(x_0 \in B) = \mu(B)$. Thus, $\mathbb{Q} = \mathbb{P}\circ\varphi_X^{-1}$ is a solution to the martingale problem $(0, \mu, \mathcal{L}_t)$. And the theorem is proved. □

Corollary 4.2.31. *There is uniqueness in law for the equation*

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t,$$

with $X_0 \sim \mu$ if and only if the martingale problem $(0, \mu, \mathcal{L}_t)$ has a unique solution.

The following theorem ensure the existence and uniqueness of a weak solution for SDEs, it was proved by D. Stroock and S.R.S. Varadhan.

Theorem 4.2.32. *Suppose $b : \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}^d \to \mathbb{R}^d$ are bounded and continuous, further, assume $a(x) = \sigma(x) \cdot \sigma(x)'$ with $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ is continuous and a is uniformly elliptic, i.e. there is a constant $C > 0$ such that, for any $z \in \mathbb{R}^d$*

$$\sum_{i,j=1}^{d} a^{ij}(t,x) z_i z_j \geqslant C|z|^2,$$

Then the martingale problem $(s, x, \mathcal{L}_t)$ has a unique solution.

Applying the transformation of the coefficients in (4.2.26) and (4.2.27), allows us to apply Theorem 4.2.32, resulting in the following theorem

Theorem 4.2.33. *Let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be continuous, bounded functions, suppose further that $a = \sigma \cdot \sigma'$ is uniformly elliptic, then the martingale problem has a unique solution for any $(s, x, \mathcal{L}_t)$.*

For proofs of these important theorems we refer the reader to [34] and [25].

2.4 Further conditions to prove tightness

In this section we adapt a general result on tightness of laws of continuous semimartingales, as an alternative to Proposition 4.2.6. While estimates of moments are very useful, the coefficients of (4.2.5) may not have a linear growth condition or boundedness that would enable us to obtain estimates such as in (4.2.4). The following results provide an alternative to Proposition 4.2.6. The general version for discontinuous semimartingales is beyond the scope of this work in general, the reader is referred to [33].
Let $\left(\Omega^n, \mathcal{F}^n, \{\mathcal{F}^n_t\}_{t\geqslant 0}, \mathbb{P}^n\right)$ be a sequence of stochastic bases satisfying the usual conditions. On each filtered probability space $\left(\Omega^n, \mathcal{F}^n, \{\mathcal{F}^n_t\}_{t\geqslant 0}, \mathbb{P}^n\right)$ suppose $\{x^n_t\}_{t\geqslant 0}$ is a continuous semimartingale (w.r.t $\left(\mathcal{F}^n_t, \mathbb{P}^n\right)$).

Assumption 4.2.1. *Let $b^n(x,w)$ be a sequence of $\mathbb{R}^d-$valued and $a^n(x,w)$ be a sequence S^d_+ valued for each n, i.e. $b^n : \mathbb{R}^+ \times \mathcal{W}^d \to \mathbb{R}^d$ and $a^n : \mathbb{R}^+ \times \mathcal{W}^d \to S^d_+$.*

a *The functions $b^n : [0,\infty) \times \mathcal{W}^d \to \mathbb{R}^d$ are locally integrable and $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{N}-$measurable and $\mathcal{N}_t-$progressively measurable.*

b *For every n, the function $a^n : [0,\infty) \times \mathcal{W}^d \to S^d_+$ is bounded and $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{N}$ measurable and $\mathcal{N}_t-$progressively measurable.*

On $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$ define the process $C^n_t(\omega) := \int_0^t a^n(s, x^n_\cdot(\omega))\, ds$.
Let us denote by $\mathbb{A} = \mathbb{R}^d \times S^d_+$, with the topology induced by the norm

$$\|(b,a)\|^2 = \sum_{i=1}^{n} b_i^2 + \sum_{i,j=1}^{n} a_{ij}^2 = \|b\|^2 + tr\left(a \cdot a'\right).$$

Assumption 4.2.2. *Suppose that for each $r \in \mathbb{R}_+$, there is a locally integrable function $L : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ such that*

I *The function $L(r,t)$ is increasing in r (i.e. in the first variable) and*

II *For all $t \in \mathbb{R}_+, y. \in \mathcal{W}^d$*

$$tr\left(a^n(t,y.)\right) + \left|b^n(t,y.)\right| \leqslant L(r,t), \tag{4.2.55}$$

whenever $\sup_{s\leqslant t}|y_s| \leqslant r$.

Under this assumptions we have the following lemma.

Lemma 4.2.34. *Suppose Assumption 4.2.1 and 4.2.2 are in force and let τ_r^n be $\mathcal{F}_t^n-$stopping times for each $n = 1,2,\dots$ and for each $r > 0$. Let $\alpha(r)$ be a real-valued finite function on $(0,\infty)$ such that*

1. *For all n and $r > 0$ we have*

$$\|x_t^n\| \leqslant \alpha(r) \ \text{ if } 0 \leqslant t < \tau_r^n. \tag{4.2.56}$$

2. *For all $T > 0$*

$$\lim_{r\to\infty} \overline{\lim}_{n\to\infty}\mathbb{P}^n\left(\tau_r^n \leqslant T\right) = 0. \tag{4.2.57}$$

Then the sequence of laws of X^n, denoted by $\mathbb{Q}^n$ is weakly relatively compact.

Define the following family of stopping times.

Definition 4.2.35. The class of stopping times $\mathcal{T}_N^n$ is given by

$$\mathcal{T}_N^n = \left\{\tau : \Omega \to \mathbb{R}_+ \ : \tau \text{ is a finite } \mathcal{F}_t^n - \text{stopping time and } \tau \geqslant N\,\mathbb{P}- \text{ a.s.}\right\}. \tag{4.2.58}$$

The following is an important characterisation of tightness of laws.

Proposition 4.2.36. *Suppose the following condition holds.*

- *For any $\epsilon > 0$ and $M > 0$*

$$\lim_{\delta\to 0}\overline{\lim_{n\to\infty}}\sup_{\tau\in\mathcal{T}_N^n}\mathbb{P}^n\left(\sup_{s\leqslant\delta}\left|X_{\tau+s}^n - X_\tau^n\right| > \epsilon\right) = 0, \tag{4.2.59}$$

- *For any $M > 0$, $\delta > 0$, $N > 0$ there is an integer $n_0(N,\delta,M)$ such that for all $n \geqslant n_0(N,\delta,M)$,*

$$\mathbb{P}^n\left(\sup_{s\leqslant N}|X_s^n| > M\right) < \delta. \tag{4.2.60}$$

Then the set of laws of $\{X_\cdot^n\}_{n\geqslant 1}$ is tight.

The conditions in Proposition 4.2.36 were proposed in [1] and are sometimes referred as the Aldous conditions, (4.2.59) and tightness of the laws of $\sup_{s<T}|X_s^n|$ in $\mathbb{R}$, imply tightness of the laws of X^n.

Remark 4.2.37. The conditions of Lemma 4.2.34 are equivalent to the following requirement: For any $T > 0$ the sequence of distributions of $\sup_{t\leqslant T}|X_t^n|$ is tight i.e.

$$\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\,\mathbb{P}^n\left(\sup_{t\leqslant T}|X_t^n| \geqslant r\right) = 0, \tag{4.2.61}$$

take $\tau_r^n = \inf\{t : |X_t^n| > r\}$, it follows that τ_r^n are $\left\{\mathcal{F}_t^{X^n}\right\}_{t\geqslant 0}$-stopping times and moreover, $\alpha(r) = r$.

Assumption 4.2.3. • *For each n there exists a non-negative $\mathcal{F}_t^n$−predictable function $L_n(t)$ such that*

$$\left\langle b^n\left(t, X_{\cdot}^n\right), X_t^n\right\rangle + tr\left(a^n\left(t, X_{\cdot}^n\right)\right) \leqslant L_n(t)\left(1 + \|X_t^n\|^2\right), \tag{4.2.62}$$

for almost all (ω, t). Furthermore

- *For any $T \in [0, \infty)$*

$$\lim_{c\to\infty} \overline{\lim_{n\to\infty}} \mathbb{P}^n\left(\int_0^T L_n(t)\, dt > c\right) = 0. \tag{4.2.63}$$

The result of interest in this section is the following theorem, the aim of studying this result is to have an alternative condition to (4.2.4) in Proposition 4.2.6, verifiable and that replaces boundedness or a linear growth condition (in diffusion models) to obtain tightness of laws.

Theorem 4.2.38. *Let X_t^n be $\mathcal{F}_t^n$- continuous semimartingales such that,*

- $\lim_{N\to\infty} \overline{\lim}_{n\to\infty} \mathbb{P}^n\left(|X_0^n| \geqslant N\right) = 0,$
- *Assumptions 4.2.1 , 4.2.2 and 4.2.3 are in force.*

Then the sequence of laws of X_t^n is weakly relatively compact.
Furthermore, let n be any integer, and $f^n : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^k$ are Borel measurable and for all (t, x) and $n \geqslant 1$

$$\left|f^n(t, x)\right| \leqslant L\left(|x|, t\right). \tag{4.2.64}$$

Define

$$Y_t^n = \int_0^t f^n\left(s, X_s^n\right) ds,$$

then the sequence of joint laws of $\left(X_s^n, Y_s^n\right)$ is tight.

Chapter 5

A SUPERMARTINGALE CHARACTERIZATION OF SETS OF SEMIMARTINGALE LAWS

1 Introduction

In this chapter we introduce an important characterization of weak limits of laws for a class of Itô processes. We introduce relevant concepts in the first section, then some important assumptions and preliminary properties related to such a characterization are presented, in the third section we describe the main result and an important consequence, namely compactness of laws for a class of Itô processes. Lastly, we discuss briefly some of the ideas behind the proofs. This chapter is a detailed account of the relevant results proved by Krylov N.V. in [32], some estimates and computations are part of our work. The results in [43] and in chapter 6 are consequences of the main results in [32] hence we deem sensible to explain the main theorems of [32] in this chapter.

The idea of approximating processes through diffusions is widely used in engineering applications, in some problems (in communications and queues, for instance) it is preferred to have a tractable problem by means of a diffusion or a Markov chain model, obtain an answer to the problem under some assumptions and then prove that under these conditions the system behaves asymptotically in this way. While in our work we do not follow this rationale, we deem worthy to recall that the theory presented is relevant in other areas and applications. In our case, as shown before in chapter 3, simple estimates yield probabilistic properties of the processes whereby we can prove existence of 'optimal' laws.

Let d be an integer, consider two measurable and adapted processes on a filtered probability space, $b_t(\omega)$ is an $\mathbb{R}^d-$valued process and $a_t(\omega)$ taking values in the set of non-negative symmetric matrices $d \times d$. Under some conditions, these two processes characterise a law on $\mathcal{B}\left(\mathcal{W}^d\right)$ by means of the corresponding martingale problem $(x, 0, \mathcal{L}_t)$, recall that a, b define the differential operator (4.2.9). Unlike the case of working with processes that are strong solutions of Itô equations (4.2.5), our assumptions does not depend on specific properties of the coefficients (i.e. Lipschitz and linear growth) but rather, they depend on properties of the set on which the mapping $(a,b) : [0,\infty) \times \Omega \to \mathbb{R}^d \times \mathbb{R}^{r\times r}$ takes values. We shall be more precise in the next sections.

2 Definitions and set-up

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $\{\mathcal{F}_t\}_{t\geqslant 0}$ be a filtration such that $\mathcal{F}_0$ is complete, $\{\mathcal{F}_t\}_{t\geqslant 0}$ is right continuous and $\mathcal{F}_t \subset \mathcal{F}$, for all t.
Let $d \geqslant 1$ be an integer and $\mathbb{A} := S_+^d \times \mathbb{R}^d$, if it is suitable, we may identify this set as a subset of $\mathbb{R}^{d_1}$ with $d_1 = d^2 + d$. Thus, we may identify an element of $\mathbb{A}$ as a d_1- dimensional vector.
For $(a,b) \in \mathbb{R}^{d_1}$, we define $\|(a,b)\|$ as the d_1-dimensional euclidean norm. Obviously, such a norm also coincides with the induced norm by the inner product $\langle\cdot,\cdot\rangle : \left(\mathbb{R}^{d\times d}\times\mathbb{R}^d\right) \times \left(\mathbb{R}^{d\times d}\times\mathbb{R}^d\right) \to \mathbb{R}$

$$\langle (a,b), (\alpha,\beta)\rangle = tr\,(a\alpha^\star) + b\cdot\beta = \sum_{i,j=1}^{d} a_{ij}\alpha_{ij} + \sum_{k=1}^{d} b_k\beta_k.$$

Thus, the norm $\|(a,b)\|$ is equal to

$$\|(a,b)\|^2 = tr\,(aa^\star) + \|b\|^2. \tag{5.2.1}$$

For each $(t,\omega) \in [0,\infty)\times\Omega$ we consider a class of bounded, closed and convex subsets $A_t(\omega) \subset \mathbb{A}$, (its bound, in general, will depend on (ω,t)). As mentioned before, instead of imposing properties on the coefficients (a,b) we shall impose conditions on the class of 'admissible' subsets $A_t(\omega)$. To each (t,ω), we define $\|A_t(\omega)\|$ as

$$\|A_t(\omega)\| := \max\left\{\|(a,b)\| \,:\, (a,b)\in A_t(\omega)\right\}. \tag{5.2.2}$$

We also define the support function of the convex set A_t, for each $(u,v)\in\mathbb{R}^{d\times d}\times\mathbb{R}^d$

$$F_t(u,v)(\omega) := \max_{(a,b)\in A_t(\omega)} \left\{tr\,(au^\star) + \langle b,v\rangle\right\},$$

here $u^\star$ denotes the transpose of u. By the properties of the subsets $A_t(\omega)$, the (random) function $F_t(u,v)$ is finite and Lipschitz continuous in (u,v). And obviously $F_t(0,0) = 0$. One can check that

$$|F_t(f,g) - F_t(u,v)| \leqslant \|A_t\|\,\|(f-u, g-v)\|, \tag{5.2.3}$$

this follows by the elementary inequality

$$\left|\sup_{x\in A} H(x) - \sup_{x\in A} G(x)\right| \leqslant \sup_{x\in A}|H(x) - G(x)|,$$

and taking $H(a,b) := tr\,(au^\star + \langle b,v\rangle)$ and $G(a,b) := tr\,(af^\star + \langle b,g\rangle)$ (G and H depend on (u,v) and (f,g) respectively). Here $G,H : \mathbb{R}^{d_1}\to\mathbb{R}$ are linear functionals and one can easily check

$$|tr\,(au^\star) + \langle b,v\rangle| \leqslant \|A_t\|\,\|(u,v)\|, \tag{5.2.4}$$

indeed, this is a consequence of Cauchy-Schwarz inequality for $tr(\cdot)$ and $|\cdot|$

$$\left|tr\left(au'\right)\right| \leqslant tr\left(aa'\right)^{1/2} tr\left(uu'\right)^{1/2} \quad \text{and} \quad |\langle b, v\rangle| \leqslant |b|\,|v|\,,$$

and applying the definition of the norm $\|A_t\|$

$$\left|tr\left(au'\right) + \langle b, v\rangle\right| \leqslant \|(a,b)\|\,\|(u,v)\| \leqslant \|A_t\left(\omega\right)\|\,\|(u,v)\|\,.$$

Notice that actually

$$\|A_t\| = \max_{\|u,v\|\leqslant 1} F_t\left(u,v\right) \tag{5.2.5}$$

Indeed, by (5.2.4)

$$F_t\left(u,v\right) \leqslant \|A_t\|\,\|(u,v)\|\,,$$

on the other hand, choosing $u^{\star}_{i,j} = \frac{a_{ij}}{\|(a,b)\|}$ and $v^{\star}_i = \frac{b_i}{\|(a,b)\|}$ yields

$$F_t\left(u^{\star}, v^{\star}\right) \geqslant \sum_{i,j=1}^{d} a_{ij}\left(\frac{a_{ij}}{\|(a,b)\|}\right) + \sum_{k=1}^{d} b_k\left(\frac{b_k}{\|(a,b)\|}\right) = \|(a,b)\|\,,$$

and $\|(u,v)\| = 1$. Thus, $\max_{\|(u,v)\|\leqslant 1} F_t\left(u,v\right) \geqslant \|A_t\|$.
The equality in (5.2.5) implies that the mapping $(t,\omega) \to \|A_t\left(\omega\right)\|$ is measurable if $F_t\left(u,v\right)$ is measurable (by taking the "sup" over a countable, dense subset of the unit ball in $\mathbb{R}^{d_1}$). In other words, the requirement of having F_t an $\mathcal{F}_t-$adapted and measurable stochastic process implies the same properties for $\|A_t\left(\omega\right)\|$, this motivates the following definition.

Definition 5.2.1. We say that the set-valued function A_t is *appropiately measurable* if $F_t\left(u,v\right)$ is $\mathcal{B}\left(\mathbb{R}^+\right)\otimes\mathcal{F}-$measurable and $\mathcal{F}_t-$adapted.

We shall make the following assumptions

Assumption 5.2.1. **1)** *The set-valued mapping $(t,\omega) \mapsto A_t\left(\omega\right)$ is appropiately measurable in the sense of definition 5.2.1.*

2) *The set-valued mapping A_t is bounded and its 'norm' $\|A_t\left(\omega\right)\|$ is locally integrable $\mathbb{P}-a.s.$ that is*

$$\int_0^T \|A_t\left(\omega\right)\|\,dt < \infty.\ \mathbb{P} - a.s. \tag{5.2.6}$$

There is an alternative way to characterise the notion of 'appropiately measurable' in Definition 5.2.1, this lemma and its proof can be found in [32].

Lemma 5.2.2. *Let $(Z, \mathcal{Z})$ be a measurable space and for each $z \in Z$, let $A\left(z\right)$ be a non-empty closed, bounded convex set in $\mathbb{R}^d$ define*

$$F\left(z,u\right) := \max_{\gamma\in A(z)} \left\langle \gamma, u\right\rangle,\ u \in \mathbb{R}^d, \tag{5.2.7}$$

let $d\left(z,a\right)$ be the distance of $A\left(z\right)$ to a point $a \in \mathbb{R}^d$. Then $A\left(z\right)$ is $\mathcal{Z}-$measurable (in the sense that $F\left(z,u\right)$ is $\mathcal{Z}-$measurable for each $u \in \mathbb{R}^d$) if and only if $d\left(z,a\right)$ is $\mathcal{Z}-$measurable for any $a \in \mathbb{R}^d$.

The condition referred in Lemma 5.2.2 seems a natural geometric condition, the family of sets A_t cannot exhibit any 'pathology'.

Remark 5.2.3. The support function F_t has the following properties

a) F_t is convex in $\mathbb{R}^d$.

b) F_t is positively homogeneous of degree 1 *w.r.t.* $(u,v) \in \mathbb{R}^{d_1}$.

c) For any $(u,v) \in \mathbb{R}^{d_1}$, $\lambda \in \mathbb{R}^d$

$$F_t\left(\frac{1}{2}\left(u+u^\star, v\right)\right) = F_t\left(u,v\right) \ , \ F_t\left(u+\lambda\cdot\lambda', v\right) \geqslant F_t\left(u,v\right).$$

Properties a) and b) in Remark 5.2.3 are straightforward as F_t is the suprema of a family of linear functionals, as A_t is bounded, previous estimations yield that F_t is finite for all $(u,v) \in \mathbb{R}^{d_1}$. The property described in c) is also a simple consequence, indeed

$$F_t\left(u+u^\star, v\right) = \max_{(a,b)}\left\{\frac{1}{2}\sum_{i,j=1}^{d}\left(u_{ij}+u_{ji}\right)a_{ij} + \sum_{i=1}^{d} b_i v_i\right\} = \max_{(a,b)}\left\{\frac{1}{2}\left(tr\left(a\cdot u^\star\right)+tr\left(u\cdot a^\star\right)\right)+\langle b, v\rangle\right\},$$

$$F_t\left(u+u^\star, v\right) = \max_{(a,b)}\left\{tr\left(a\cdot u^\star\right) + \sum_{i=1}^{d} b_i v_i\right\}.$$

As $tr\left(B\right) = tr\left(B^\star\right)$ and a is symmetric. Similarly, as a is non-negative definite.

$$F_t\left(u+\lambda\cdot\lambda', v\right) = \max_{(a,b)\in A_t(\omega)}\left\{\sum_{i,j=1}^{d}\left(u_{ij}+\lambda_i\lambda_j\right)a_{ij} + \sum_{i=1}^{d} v_i b_i\right\} \geqslant \max_{(a,b)\in A_t(\omega)}\left\{\sum_{i,j=1}^{d} u_{ij}a_{ij} + \sum_{i=1}^{d} v_i b_i\right\}.$$

It is possible to start with a family of support functions satisfying the properties a)-c) in Remark 5.2.3 and define the family of convex sets in terms of F_t.

$$A_t = \left\{(a,b)\in\mathbb{R}^{d_1} : \sum_{i,j=1}^{d} a_{ij}u_{ij} + \sum_{i=1}^{d} b_i v_i \leqslant F_t\left(u,v\right)\right\} = \left\{(a,b)\in\mathbb{R}^{d_1} : \langle (a,b),(u,v)\rangle_{\mathbb{R}^{d_1}} \leqslant F_t\left(u,v\right)\right\}. \tag{5.2.8}$$

For further details the reader is referred to [32].

3 Main theorem

We describe the main theorem of this chapter, this result is closely related to Theorems 4.2.29 and 4.2.30 in chapter 4. Furthermore, the main assumption resembles the definition of a martingale problem, Definition 4.2.18 and (4.2.21) which in turn describes the existence of a weak solution of a Itô SDE. See Theorem 2.2 in [32].
Recall that $C_0^\infty\left(\mathbb{R}^d;\mathbb{R}\right)$ denotes the space of smooth functions with compact support.

Main Theorem 5.3.1. *Suppose* $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ *is a filtered probability space such that* $\mathcal{F}_0$ *is complete and the filtration satisfies the usual hypothesis.*
Let Assumption 5.2.1 be in force and let X_t *be continuous* $\mathcal{F}_t-$*adapted,* $\mathbb{R}^d$*-valued process such that for each* $u \in C_0^\infty\left(\mathbb{R}^d;\mathbb{R}\right)$ *the process*

$$\eta_t\left(u\right) := u\left(X_t\right) - \int_0^t F_s\left(u_{xx}\left(X_s\right), u_x\left(X_s\right)\right)ds \tag{5.3.1}$$

is a local $\mathcal{F}_t-$supermartingale.
Then, on $\Omega \times [0,\infty)$ there exists an $\mathbb{A}-$valued process (a_t, b_t) such that

i) *the process (a_t, b_t) is $\mathcal{B}([0,\infty)) \otimes \mathcal{F}$-measurable and $\mathcal{F}_t-$adapted.*

ii) *for almost all $\omega \in \Omega$ and $t \geqslant 0$ it holds that $(a_t, b_t) \in A_t$*

iii) *there exists an extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ (in the sense of Definition 4.2.25) and a d-dimensional Wiener process defined on this extension, W_t, such that for every $r \in [0,\infty)$, $\{X_s, s \in [0,r]\}$ and $W_t - W_r$ are independent and with probability one*

$$X_t = X_0 + \int_0^t \sqrt{2a_s}dW_s + \int_0^t b_s ds \ \ \textit{for all}\ t \geqslant 0. \tag{5.3.2}$$

Notice that we do not make a distinction (in terms of notation) between the processes defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and denote by $X_s(\omega)$, $a_s(\omega)$ and $b_s(\omega)$ their respective extended processes on $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left\{\tilde{\mathcal{F}}_t\right\}_{t\geqslant 0}, \tilde{\mathbb{P}}\right)$.
Recalling the estimates in (5.2.3) applied to $(f,g) = (0,0)$ we have

$$|F_t(u_{xx}(X_s), u_x(X_s))| \leqslant \|A_t\| \, \|(u_{xx}, u_x)\| \leqslant N \cdot \|A_t\|,$$

since the function u_{xx}, u_x and u are bounded, we have

$$|F_s(u_{xx}(X_s), u_x(X_s))| \leqslant M \cdot \|A_t\|,$$

and the constant does not depend on X_s but only on u, u_x and u_{xx}.
If A_t is not a convex set (or at least it is not connected), Theorem 5.3.1 may not hold. Indeed, suppose $A_t = \left\{(a,b) \in \mathbb{R}^d : \ (a,b) = (0,0), (0,2)\right\}$ taking $X_t = t$ we have that the assumption holds, for any $u \in C_0^\infty\left(\mathbb{R}^d; \mathbb{R}\right)$

$$u(X_t) - u(X_s) = \int_s^t u'(X_r)\, dr \leqslant \int_s^t \max\left(2 \cdot u'(X_r), 0 \cdot u'(X_r)\right) dr = \int_s^t \left(2 \cdot u'(X_r)\right)_+ dr$$

this means $\eta_t(u)$ is a local supermartingale and $X_t = x + \int_0^t 1ds + \int_0^t 0 \cdot dW_s$ but $(0,1) \notin A_t$. Thus, the convexity assumption may be relaxed or replaced, but in general the theorem does not hold without an additional structure on the family of sets $A_t(\omega)$.

Remark 5.3.2. Although the theorem is stated for processes defined for all $t \geqslant 0$. One can clearly obtain a version of the theorem for processes defined on the finite interval $[0,T]$ by applying Theorem 5.3.1 to the stopped process x^T and defining $A_t = (0,0)$ for $t \geqslant T$.

4 Consequences of Theorem 5.3.1

In this section some consequences of Theorem 5.3.1 are described, we shall see that one of the important corollaries is that under some assumptions the set of laws of Itô processes is weakly closed. This in turn will allow us to prove that a 'direct method' argument is possible to apply to a family of stochastic control problems, whenever one considers a weak-setting of the dynamics.

As in previous sections, we consider the topological space $(\mathcal{W}^d, \tau_{uc})$ with τ_{uc} the topology generated by the uniform convergence on bounded subsets of $[0,\infty)$, and then $(\mathcal{W}^d, \mathcal{N})$ is a measurable space, recall that the law of a stochastic process is a probability measures on $(\mathcal{W}^d, \mathcal{N})$ and consider the filtration $\mathcal{N}_t = \sigma\{\alpha_s : s \leqslant t\} = \sigma\left(\{x. : x_s \in \Gamma\} : \Gamma \in \mathcal{B}\left(\mathbb{R}^d\right), s \leqslant t\right)$ where α_t is the truncation operator in Definition 4.2.2, and adapt the setting that was described in the previous section on this filtered probability space.
Consider a class of closed convex subsets of $\mathbb{A}$, $\{A_t(x.); t \geqslant 0, x. \in \mathcal{W}^d\}$ and its associated support function

$$F_t(u,v,x.) = \sup\left\{\sum_{i,j=1}^d a_{ij}u_{ij} + \sum_{i=1}^d b_i v_i \ : \ (a,b) \in A_t(x.)\right\}. \tag{5.4.1}$$

We shall make the following assumptions on the family of convex sets and their support functions

Assumption 5.4.1. **i)** *There exists a constant $K > 0$ such that*

$$\|A_t(x.)\| \leqslant K(1 + |x_t|) \text{ for all } t \in [0,\infty), \ x \in \mathcal{W}^d. \tag{5.4.2}$$

ii) *For all $(u,v) \in \mathbb{R}^{d_1}$ (fixed), $F_t(u,v,x.)$ is $\mathcal{B}([0,\infty)) \otimes \mathcal{N}-$measurable and for each t, $t \geqslant 0$, and (u,v) fixed, the support function $F_t(u,v,x.)$ is $\mathcal{N}_t-$adapted.*

iii) *The set-valued function $x. \to A_t(x.)$ is upper semicontinuous w.r.t to $x.$ for each $(u,v) \in \mathbb{R}^{d_1}$ and $t \in [0,\infty)$ in the sense that for each $(u,v) \in \mathbb{R}^{d_1}$ and $t \in [0,\infty)$*

$$\limsup_{(y-x)_t^* \to 0} F_t(u,v,x.) \leqslant F_t(u,v,x.) \text{ for all } x \in \mathcal{W}^d. \tag{5.4.3}$$

Let us denote by π the set consisting of a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$, whose filtration is complete w.r.t $\mathbb{P}$, and two $\mathcal{F}_t$-adapted continuous processes $\{X_t\}_{t\geqslant 0}$ and $\{W_t\}_{t\geqslant 0}$ with W_t is an $\mathcal{F}_t$-Wiener process. Thus, $\pi := (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, \{X_t\}_{t\geqslant 0}, \{W_t\}_{t\geqslant 0})$.

Definition 5.4.1. We define the class Π to be the set of all $\pi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ such that there exist an $\mathbb{A}$-valued, $\mathcal{B}([0,\infty)) \otimes \mathcal{F}-$measurable and $\mathcal{F}_t$-adapted process (a_t, b_t) such that

i $\mathbb{P} - a.s.$ for all $t \in [0,\infty)$

$$X_t = \int_0^t \sqrt{2a_s}\, dW_s + \int_0^t b_s ds, \tag{5.4.4}$$

ii For almost all $(\omega, t) \in \Omega \times [0,\infty)$, we have $(a_t, b_t) \in A_t(X.)$.

We could refer to Π as the set of admissible 'stochastic bases'. We indicate with a superscript $\Omega = \Omega^\pi$, $\mathcal{F} = \mathcal{F}^\pi$, ..., $W_t = W_t^\pi$ to denote that such an element belongs to the same π.
We recall the following well-known lemma

Lemma 5.4.2. *Let (X,d) be a metric space and $\mu_1, \mu_2, \ldots$ be a probability measures on (X,d). Assume that $\mu_n \Longrightarrow \mu$ weakly. Let $F(x)$ be a bounded upper-semicontinuous function on x*

$$\overline{\lim_{y\to x}} F(y) \leqslant F(x) \text{ for all } x \in X. \tag{5.4.5}$$

Then

$$\overline{\lim_{n\to\infty}} \int_X F(x)\, \mu_n(dx) \leqslant \int_X F(x)\, \mu(dx), \tag{5.4.6}$$

Proof. Without loss of generality we can assume that $0 \leqslant F \leqslant 1$, by hypothesis on the u.s.c. of F, $\{x \in X : F(x) \geqslant c\}$ is closed for all $c \in \mathbb{R}$, then by Fubini's theorem,

$$\int_X F(x)\,\mu_n(dx) = \int_0^1 \mu_n(F(x) \geqslant c)\,dc,$$

Fatou's lemma implies

$$\overline{\lim_n} \int_X F(x)\,\mu_n(dx) = \overline{\lim_n} \int_0^1 \mu_n(F(x) \geqslant c)\,dc \leqslant \int_0^1 \overline{\lim_n} \mu_n(F(x) \geqslant c)\,dc.$$

And by portmanteau lemma, Proposition 2.1.7

$$\int_0^1 \overline{\lim_n} \mu_n(F(x) \geqslant c)\,dc \leqslant \int_0^1 \mu(F(x) \geqslant c)\,dc = \int_X F(x)\,\mu(dx).$$

□

The following corollary of Theorem 5.3.1 is fundamental for the forthcoming sections.

Theorem 5.4.3. *Denote by $\mathbb{Q}^\pi$ the distribution of $x.$ on $\mathcal{W}^d$. Then, the set $\{Q^\pi : \pi \in \Pi\}$ is convex and sequentially weakly compact i.e. for any sequence of 'stochastic bases' (i.e. filtered probability spaces with processes (X, B) satisfying (5.4.4) defined on it), $\{\pi_n\}\ n \geqslant 1 \in \Pi$ there is a subsequence $n_m \to \infty$ and a $\pi \in \Pi$ such that for any real-valued, bounded continuous function $H(x.) : \mathcal{W}^d \to \mathbb{R}$*

$$\lim_{m\to\infty} \mathbf{E}^{\pi_{n_m}} H\left(x_\cdot^{\pi_{n_m}}\right) = \mathbf{E}^\pi H(x_\cdot^\pi). \tag{5.4.7}$$

Proof. One can check that under our assumptions, the following holds $\sup_{\pi\in\Pi} \mathbf{E}^\pi\left(\sup_{t\leqslant T} |x_t^\pi|^2\right) < \infty$. Indeed, by a localisation argument, it is possible to assume that x_t is bounded on $t \leqslant T$ and monotone convergence would give the required moment estimate for the general case. It is enough to see that $\mathbf{E}^\pi\left(\sup_{t\leqslant T} |x_t^\pi|^2\right) < \infty$ and does not depend on π .
On $\left(\Omega^\pi, \mathcal{F}^\pi, \{\mathcal{F}_t^\pi\}_{t\geqslant 0}, \mathbb{P}^\pi\right)$, we have the usual estimate,

$$\mathbf{E}^\pi\left[\sup_{t\leqslant T} |x_t|^2\right] \leqslant 2\left(\mathbf{E}^\pi\left(\int_0^T |b_s|\,ds\right)^2 + \mathbf{E}^\pi\left[\sup_{t\leqslant T}\left|\int_0^t \sqrt{2a_s}\,dw_s\right|^2\right]\right), \tag{5.4.8}$$

Applying Burkholder-Davis-Gundy inequality, see [47]

$$\mathbf{E}^\pi\left[\sup_{0\leqslant t\leqslant T} |x_t|^2\right] \leqslant 2\left(\mathbf{E}^\pi\left(\int_0^T |b_s|\,ds\right)^2 + 4\mathbf{E}^\pi\left[\int_0^T tr(2a_s)\,ds\right]\right),$$

By Assumption 5.4.1, $\left(|b_s|^2 + \|a_s\|^2\right)^{1/2} \leqslant K(1 + |x_s|)$, where $\|a_s\| = tr(a_s a_s')^{1/2}$. For simplicity, we have omitted the superscript π, by this inequality and by Cauchy-Schwarz inequality we can estimate

$$\mathbf{E}^\pi\left(\int_0^T |b_s|\,ds\right)^2 \leqslant \mathbf{E}^\pi\left(\int_0^T \|A_s\|\,ds\right)^2 \leqslant 2K^2 T\mathbf{E}^\pi \int_0^T 1 + |x_s|^2\,ds,$$

by Cauchy-Schwarz inequality applied to $\langle A, B\rangle = tr(AB')$ (defined on $\mathbb{R}^{d\times d}$)

$$\mathbf{E}^\pi \int_0^T tr(2a_s)\,ds \leqslant 2\sqrt{d}\mathbf{E}^\pi \int_0^T tr(a_s a_s')^{\frac{1}{2}}\,ds \leqslant 2\sqrt{d}K\mathbf{E}^\pi \int_0^T 1+|x_s|\,ds \leqslant 4\sqrt{d}K\mathbf{E}^\pi \int_0^T 1+|x_s|^2\,ds,$$

since $(1+|x|)^2 \leqslant 2\left(1+|x|^2\right)$, then

$$\mathbf{E}^\pi\left[\sup_{t\leqslant T}|x_t|^2\right] \leqslant 2\left(2K^2T^2 + 2K^2T\mathbf{E}^\pi\int_0^T |x_s|^2\,ds + 4\sqrt{d}KT + 4\sqrt{d}K\mathbf{E}^\pi\int_0^T |x_s|^2\,ds\right), \tag{5.4.9}$$

$$\mathbf{E}^\pi\left[\sup_{t\leqslant T}|x_t|^2\right] \leqslant 4K^2T^2 + 8\sqrt{d}KT + \left(4K^2T + 8\sqrt{d}K\right)\mathbf{E}^\pi\int_0^T \sup_{t\leqslant s}|x_t|^2\,ds. \tag{5.4.10}$$

We can use an argument similar to the one that proves Grömwall inequality to verify that (5.4.9) implies

$$\mathbf{E}^\pi\left[\sup_{s\leqslant T}|x_s|^2\right] \leqslant g(T) + \left(8K\sqrt{d} + 4K^2T\right)\int_0^T e^{\int_s^T 8K\sqrt{d}+4K^2r dr} g(s)\,ds, \tag{5.4.11}$$

with $g(t) = 4K^2t^2 + 8K\sqrt{d}t$. In other words, as we had supposed that $\sup_{s\leqslant T}|x_s|^2$ was bounded, then (5.4.11) applies to $\sup_{s\leqslant t\wedge\tau_m}|x_s|^2$ with $\tau_m := \inf\{t>0 : |x_t| > m\}$ and monotone convergence (letting $m\to\infty$) implies that (5.4.11) holds, without this restriction.
Furthermore, the function $g(t)$ and the constants in (5.4.11) do not depend on π. Moreover, this idea can be applied to $2+\alpha$ with $\alpha>0$ to obtain similar moment estimates. In this case, we can estimate

$$\mathbf{E}^\pi\left[\sup_{0\leqslant t\leqslant T}|x_t|^{2+\alpha}\right] \leqslant 2^{\alpha+1}\left(\mathbf{E}^\pi\Big(\int_0^T |b_s|\,ds\Big)^{2+\alpha} + 2^{1+\frac{\alpha}{2}}c_\alpha d^{\frac{1}{2}+\frac{\alpha}{4}}\mathbf{E}^\pi\left[\int_0^T tr\left(a_s a_s'\right)^{1/2}ds\right]^{1+\frac{\alpha}{2}}\right),$$

$$\mathbf{E}^\pi\Big(\int_0^T |b_s|\,ds\Big)^{2+\alpha} \leqslant T^{\alpha+1}\mathbf{E}^\pi\Big(\int_0^T \|A_s\|^{2+\alpha}ds\Big) \leqslant T^{\alpha+1}K^{2+\alpha}2^{\alpha+1}\mathbf{E}^\pi\Big[\int_0^T (1+|x_s|^{2+\alpha})ds\Big]$$

$$\mathbf{E}^\pi\left(\sup_{s\leqslant T}\left|\int_0^s \sqrt{2a_s}dw_s\right|^{2+\alpha}\right) \leqslant c_\alpha\mathbf{E}^\pi\Big(\int_0^T tr(2a_r)dr\Big)^{1+\frac{\alpha}{2}} \leqslant c_\alpha 2^{1+\frac{\alpha}{2}}d^{\frac{1}{2}+\frac{\alpha}{4}}\mathbf{E}^\pi\Big(\int_0^T tr(a_r a_r')^{1/2}dr\Big)^{1+\frac{\alpha}{2}},$$

$$\mathbf{E}^\pi\Big(\sup_{s\leqslant T}\left|\int_0^s \sqrt{2a_r}dw_r\right|^{2+\alpha}\Big) \leqslant 2^{1+\frac{\alpha}{2}}d^{\frac{1}{2}+\frac{\alpha}{4}}K^{1+\frac{\alpha}{2}}c_\alpha T^{\frac{\alpha}{2}}\mathbf{E}^\pi\left[\int_0^T \left(1+|x_s|\right)^{2+\alpha}ds\right]$$

as $\|A_s\| \leqslant K\big(1+|x_s|\big)$. Applying the same arguments (as in the proof of Grömwall inequality), we obtain the following moment bound for $\sup_{t\leqslant T}|x_s|^{2+\alpha}$

$$\mathbf{E}^\pi\left[\sup_{s\leqslant T}|x_s|^{2+\alpha}\right] \leqslant C(\alpha,K)T\cdot h(T) + C(\alpha,K)\int_0^T \exp\Big(\int_t^T h(r)dr\Big)th(t)dt \tag{5.4.12}$$

where $h(t) = t^{\alpha+1} + t^{\frac{\alpha}{2}}$ and with $C(\alpha,K)$ a constant depending on d, K and α. This is an inequality similar to (5.4.11) and the constant $C(\alpha,K)$ and the function $h(t)$ does not depend on x. As it was explained before, by localisation we can assume first that x_s is bounded, and monotone convergence yields the conclusion in the general case.
These inequalities allow, in turn to estimate $\mathbf{E}^\pi|x_t-x_s|^{2+\alpha}$, let $t,s\leqslant T$. Indeed,

$$\mathbf{E}^\pi|x_t-x_s|^{2+\alpha} \leqslant 2^{2+\alpha}\mathbf{E}^\pi\left[\Big(\int_s^t |b_r|\,dr\Big)^{2+\alpha}\right] + 2^{2+\alpha}\mathbf{E}^\pi\left[\Big|\int_s^t \sqrt{2a_r}dw_r\Big|^{2+\alpha}\right],$$

by moments estimates for stochastic integrals (see Corollary 2.5.3 in [36]) (or by Burkholder-

Davis-Gundy inequality) and by Hölder inequality (Cauchy-Schwarz inequality in the second term)

$$\mathbf{E}^{\pi}\left|x_t - x_s\right|^{2+\alpha} \leqslant 2^{2+\alpha}K^{2+\alpha}\mathbf{E}^{\pi}\left(\int_s^t 1 + |x_r|\, dr\right)^{2+\alpha} + c_\alpha |t-s|^{\frac{\alpha}{2}}\mathbf{E}^{\pi}\Big(\int_s^t (tr\,(2a_r))^{1+\frac{\alpha}{2}} dr\Big), \tag{5.4.13}$$

And by assumption 5.4.1

$$\leqslant 2^{2+\alpha}K^{2+\alpha}\left|t-s\right|^{1+\alpha}\mathbf{E}^{\pi}\int_s^t \left(1+|x_r|\right)^{2+\alpha} dr + 2^{1+\alpha}c_\alpha |t-s|^{\alpha/2}K^{2+\alpha}\mathbf{E}^{\pi}\int_s^t \big(1+|x_r|\big)^{2+\alpha} dr, \tag{5.4.14}$$

applying again, Hölder inequality to the integral

$$\left(\int_s^t 1 + |x_r|\, dr\right)^{1+\beta} \leqslant (t-s)^{\beta}\int_s^t \left(1+|x_r|\right)^{1+\beta} dr \leqslant 2^{\beta}\left(t-s\right)^{\beta}\left(t-s+(t-s)\sup_{r\leqslant t}|x_r|^{1+\beta}\right).$$

Applying this estimates to $\beta = 1+\alpha$ and to $\beta = \frac{\alpha}{2}$,

$$\mathbf{E}^{\pi}\left|x_t - x_s\right|^{2+\alpha} \leqslant 2^{2+2\alpha}K^{1+\alpha}\left|t-s\right|^{2+\alpha}\left(1 + \mathbf{E}^{\pi}\sup_{s\leqslant r\leqslant t}|x_r|^{2+\alpha}\right) + \tag{5.4.15}$$

$$2^{1+\alpha}c_\alpha K^{2+\alpha}\left|t-s\right|^{\frac{\alpha}{2}}\mathbf{E}^{\pi}\int_s^t \big(1+\sup_{u\leqslant r}|x_u|\big)^{2+\alpha} dr,$$

$$\mathbf{E}^{\pi}\left|x_t - x_s\right|^{2+\alpha} \leqslant 2^{2+2\alpha}K^{1+\alpha}\left|t-s\right|^{2+\alpha}\left(1 + \mathbf{E}^{\pi}\sup_{r\leqslant T}|x_r|^{2+\alpha}\right) + \tag{5.4.16}$$

$$2^{1+\alpha}c_\alpha K^{2+\alpha}\left|t-s\right|^{1+\frac{\alpha}{2}}\left(1 + \mathbf{E}^{\pi}\sup_{r\leqslant T}|x_r|^{2+\alpha}\right). \tag{5.4.17}$$

By (5.4.12) we obtain the inequality

$$\mathbf{E}^{\pi}\left|x_t - x_s\right|^{2+\alpha} \leqslant \mathcal{N}\left(K,\alpha,d,T\right)\left|t-s\right|^{1+\frac{\alpha}{2}}. \tag{5.4.18}$$

Denote by $\mathbb{Q}^{\pi}$ the law of the process x_t^{π}, a consequence of the moment bound (5.4.18) and (5.4.11) is the tightness of the family of probability measures $\{\mathbb{Q}^{\pi} : \pi \in \Pi\}$, see Proposition 4.2.6. Thus, if $\{\mathbb{Q}^{\pi_n} : \pi_n \in \Pi\}$ is a sequence of laws, of processes $x_t^n \in \pi_n$, by Prokhorov's theorem there must be a subsequence and a probability measure such that $\mathbb{Q}^{n_k} \Rightarrow \nu$, with ν a probability measure on $\mathcal{W}^d$.

Let $x_t : \mathcal{W}^d \to \mathbb{R}$ be the canonical process on the probability space $\left(\mathcal{W}^d, \mathcal{N}, \mathcal{N}_t, \nu\right)$, ν is the law of the process x_t. In order to prove that $\{\mathbb{Q}^{\pi} : \pi \in \Pi\}$ is weakly compact, there must be a filtered probability space $\left(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathcal{F}_t}, \overline{\mathbb{Q}}\right)$ and a $\overline{\mathcal{F}_t}-$ Wiener process, $\{\overline{w}_t\}$ such that

$$X^{\pi}_{\cdot} := x_{\cdot}, \tag{5.4.19}$$

and (5.4.4) holds in other words, x_t is the canonical process and (5.4.4) holds on the probability space $\left(\mathcal{W}^d, \mathcal{B}\left(\mathcal{W}^d\right), \nu\right)$. For each $\pi \in \Pi$, $u \in C_0^{\infty}\left(\mathbb{R}^d; \mathbb{R}\right)$ and any set of points $0 \leqslant t_1 \leqslant \ldots \leqslant$

$t_q \leqslant s \leqslant t < \infty$ and non-negative continuous bounded function f on $\mathbb{R}^{qd}$,

$$\mathbf{E}^\pi\left[f\left(x^\pi_{t_1},\ldots,x^\pi_{t_q}\right)\left(u\left(x^\pi_t\right)-u\left(x^\pi_s\right)\right)\right]=\mathbf{E}^\pi\left[f\left(x^\pi_{t_1},\ldots,x^\pi_{t_q}\right)\int_s^t\left\{\sum_{i,j=1}^d a^\pi_{ij}(r)u_{x_ix_j}\left(x^\pi_r\right)+\ldots\right.\right.$$
$$\left.\left.\sum_{i=1}^d b^\pi_i(r)u_{x_i}\left(x^\pi_r\right)\right\}dr+\int_s^t f\left(x^\pi_{t_1},\ldots,x^\pi_{t_q}\right)\sum_{i=1}^d\sum_{k=1}^d u_{x_i}\left(x^\pi_r\right)\sqrt{2a^\pi_{i,k}}(r)dw^\pi_k(r)\right]. \tag{5.4.20}$$

Thus,

$$\mathbf{E}^\pi\left[f\left(x^\pi_{t_1},\ldots,x^\pi_{t_q}\right)\left(u\left(x^\pi_t\right)-u\left(x^\pi_s\right)\right)\right]\leqslant\mathbf{E}^\pi\left[f\left(x^\pi_{t_1},\ldots,x^\pi_{t_q}\right)\int_s^t F_r\left(u_{x_ix_j}\left(x^\pi_r\right),u_{x_i}\left(x^\pi_r\right),x^\pi_\cdot\right)dr\right]. \tag{5.4.21}$$

From the inequality (5.4.21) applied to the subsequences $\{\pi_{n_m}\}_{m\geqslant 1}$ and letting $m\to\infty$ then

$$\int_{\mathcal{W}^d} f\left(p_{t_1}(x_\cdot),\ldots,p_{t_q}(x_\cdot)\right)\left(u\left(p_t\right)-u\left(p_s\right)\right)d\nu\left(x_\cdot\right)\leqslant\overline{\lim_{m\to\infty}}\mathbf{E}^{\pi_{n_m}}\left[f\left(x^{\pi_{n_m}}_{t_1},\ldots,x^{\pi_{n_m}}_{t_q}\right)\cdot\right.$$
$$\left.\int_s^t F_r\left(u_{x_ix_j}\left(x^{\pi_{n_m}}_r\right),u_{x_i}\left(x^{\pi_{n_m}}_r\right),x^\pi_\cdot\right)dr\right]. \tag{5.4.22}$$

The support function $(r,y_\cdot)\to F_r\left(u_{x_ix_j}\left(y_r\right),u_{x_i}\left(y_r\right),y_\cdot\right)$ is bounded.
First of all, by Assumption 5.4.1 and (5.2.5) we have

$$\sup_{|(u,v)|\leqslant 1} F_t\left(u,v,x_\cdot\right)\leqslant K\left(1+|x_t|\right).$$

If $\rho=\inf\left\{M>0\,:\,\{supp\, u\},\{supp\, u_{x_i}\},\left\{supp\, u_{x_i,x_j}\,:i,j\right\}\subset[-M,M]\right\}$ and if $|y_r|\leqslant\rho$ then by homogeneity of the support function

$$F_t\left(u_{x_ix_j}\left(y_r\right),u_{x_i}\left(y_r\right),y^\pi_\cdot\right)=\rho\cdot F_t\left(\frac{u_{x_ix_j}}{\rho}\left(y_r\right),\frac{u_{x_i}}{\rho}\left(y_r\right),y^\pi_\cdot\right)\leqslant\rho\left\|A_t\left(y_\cdot\right)\right\|. \tag{5.4.23}$$

And by Assumption 5.4.1,

$$\rho\left\|A_t\left(y_\cdot\right)\right\|\leqslant\rho K\cdot\left(1+|y_t|\right)\leqslant K\rho\left(1+\rho\right).$$

Notice that ρ depends only on the function u, and this is enough, since in case $|y_r|\geqslant\rho$, the support function $F_t\left(u_{x_ix_j}\left(y_t\right),u_{x_i}\left(y_t\right),y_\cdot\right)=0$. Furthermore, fixing $u\in C_0^\infty\left(\mathbb{R}^d;\mathbb{R}\right)$ the function

$$y_\cdot\to f\left(y_{t_1},\ldots,y_{t_q}\right)\int_s^t F_r\left(u_{x_ix_j}\left(y_r\right),u_{x_i}\left(y_r\right),y_\cdot\right)dr$$

is upper-semicontinuous.
Fix $t<\infty$ and suppose $y^n\to x$ in $\mathcal{W}^d$ for all $t>0$ by Fatou's lemma, the fact that the function is bounded and Assumption 5.4.1.

$$\overline{\lim_{(y^n-x)^\star_t\to 0}} f\left(y^n_{t_1},\ldots,y^n_{t_q}\right)\int_s^t F_r\left(u_{x_ix_j}\left(y^n_r\right),u_{x_i}\left(y^n_r\right),y^n_\cdot\right)dr\leqslant f\left(x_{t_1},\ldots,x_{t_q}\right)\cdot$$
$$\int_s^t\overline{\lim_{(y^n-x)^\star_t\to 0}}F_r\left(u_{x_ix_j}\left(y^n_r\right),u_{x_i}\left(y^n_r\right),y^n_\cdot\right)dr. \tag{5.4.24}$$

By decomposing the integrand in the right hand side of (5.4.24), one can check that by Assumption 5.4.1 and Lemma 5.4.2

$$\begin{aligned} F_r\left(u_{xx}\left(y_r^n\right), u_x\left(y_r^n\right), y_\cdot^n\right) &= F_r\left(u_{xx}\left(y_r^n\right), u_x\left(y_r^n\right), y_\cdot^n\right) - F_r\left(u_{xx}\left(y_r^n\right), u_x\left(y_r^n\right), x_\cdot\right) + \\ F_r\left(u_{xx}\left(y_r^n\right), u_x\left(y_r^n\right), x_\cdot\right) &- F_r\left(u_{xx}\left(x_r\right), u_x\left(x_r\right), x_\cdot\right) + F_r\left(u_{xx}\left(x_r\right), u_x\left(x_r\right), x_\cdot\right). \end{aligned} \tag{5.4.25}$$

As $n \to \infty$ we obtain

$$\overline{\lim_{n\to\infty}} F_r\left(u_{xx}\left(y_r^n\right), u_x\left(y_r^n\right), y_\cdot^n\right) \leqslant F_r\left(u_{xx}\left(x_r\right), u_x\left(x_r\right), x_\cdot\right). \tag{5.4.26}$$

Hence from (5.4.26) and (5.4.24) we have

$$\underline{\lim}_{(y^n - x)_t^\star \to 0} f\left(y_{t_1}^n, \ldots, y_{t_q}^n\right) \int_s^t F_r\left(u_{x_i x_j}\left(y_r^n\right), u_{x_i}\left(y_r^n\right), y_\cdot^n\right) dr \leqslant f\left(x_{t_1}, \ldots, x_{t_q}\right) \tag{5.4.27}$$

$$\int_s^t F_r\left(u_{x_i x_j}\left(x_r\right), u_{x_i}\left(x_r\right), x_\cdot\right) dr, \tag{5.4.28}$$

and the claim follows.
By Lemma 5.4.2 and the fact $\mathbb{Q}^{\pi_{n_m}} \Rightarrow \nu$,

$$\begin{aligned} \mathbf{E}^\nu\left[f(p_{t_1}, \ldots, p_{t_q})(u\left(p_t\right) - u\left(p_s\right))\right] &\leqslant \overline{\lim_{m\to\infty}} \mathbf{E}^{n_m}\left[f\left(x_{t_1}^{n_m}, \ldots, x_{t_q}^{n_m}\right) \int_s^t F_r\left(u_{xx}\left(x_r^{n_m}\right), u_x\left(x_r^{n_m}\right), x_\cdot^{n_m}\right) dr\right] \\ &\leqslant \mathbf{E}^\nu\left[f\left(p_{t_1}, \ldots, p_{t_q}\right) \int_s^t F_r\left(u_{xx}\left(p_r\right), u_x\left(p_r\right), p_\cdot\right) dr\right]. \end{aligned} \tag{5.4.29}$$

Hence, taking limits in (5.4.21) yields

$$\mathbf{E}^\nu\left[f\left(p_{t_1}, \ldots, p_{t_q}\right)\Big(u\left(p_t\right) - u\left(p_s\right)\Big)\right] \leqslant \mathbf{E}^\nu\left[f\left(p_{t_1}, \ldots, p_{t_q}\right) \int_s^t F_r\left(u_{xx}\left(p_r\right), u_x\left(p_r\right), p_\cdot\right) dr\right]. \tag{5.4.30}$$

Taking $A_t(\omega) := A_t(x_\cdot(\omega))$ (5.4.30) implies that

$$\eta_t(u) := u\left(p_t\right) - \int_0^t F_s\left(u_{xx}\left(p_s\right), u_x\left(p_s\right), p_\cdot\right) ds$$

is and $\mathcal{N}_t$-supermartingale, as $F_t(u, v, x_\cdot)$ is u.s.c. the set-valued mapping $A_t(x_\cdot)$ is 'appropiately measurable' (see Assumption 5.2.1). The existence of $\overline{\pi}$ satisfying (5.4.19) follows directly from Theorem 5.3.1. Thus the set $\{\mathbb{Q}^\pi : \pi \in \Pi\}$ is weakly compact.
By Theorem 5.3.1 the law $\mathbb{Q}$ belongs to $\{\mathbb{Q}^\pi : \pi \in \Pi\}$ if and only if

$$\int_{\mathcal{W}^d} f\left(x_{t_1}, \ldots, x_{t_q}\right)\left[u\left(x_t\right) - u\left(x_s\right)\right] \mathbb{Q}(dx) \leqslant \int_{\mathbb{W}^d} f\left(x_{t_1}, \ldots, x_{t_q}\right) \cdot \tag{5.4.31}$$

$$\int_s^t F_r\left(u_{xx}\left(x_r\right), u_x\left(x_r\right), x_\cdot\right) dr \mathbb{Q}(dx) \tag{5.4.32}$$

for all $q \geqslant 1$, $0 \leqslant t_i \leqslant s \leqslant t$, $u \in C_0^\infty\left(\mathbb{R}^d; \mathbb{R}\right)$ and non negative continuous bounded functions f. Both sides of (5.4.31) are linear with respect to $\mathbb{Q}$ and the condition (5.4.31) holds for linear combinations of elements in $\{\mathbb{Q}^\pi : \pi \in \Pi\}$. The theorem is proved. □

Notice that in the case of having the class of subsets $A_t(x_\cdot)$ not depending on $t \in [0, \infty)$, $y_\cdot \in$

$\mathcal{W}^d$, then the support function $F_t(u, v, y.)$ does not depend on $r, y.$ and is (jointly) measurable in this variables. And in this case, we have the following corollary

Corollary 5.4.4. *Let A be a convex, closed bounded subset of $\mathbb{A}$ and assume that for each $n = 1, 2, \ldots$ we are given an Ito process*

$$x_t^n = x_0^n + \int_0^t \sqrt{2a_s^n} dw_s^n + \int_0^t b_s^n ds, \tag{5.4.33}$$

defined on a probability space perhaps depending on n carrying a $d-$dimensional Wiener process w_t^n and appropiately measurable $A-$valued processes (a_t^n, b_t^n). Denote by $\mathbb{Q}^n$ the distribution of $x_\cdot^n$ on $\mathcal{W}^d$. Then there is a subsequence $n(k) \to \infty$ such that $\mathbb{Q}^{n(k)}$ converges weakly on $\mathcal{W}^d$ to the distribution of an Ito process

$$x_t = x_0 + \int_0^t \sqrt{2a_s} dw_s + \int_0^t b_s ds, \tag{5.4.34}$$

defined on a probability space carrying a $d-$dimensional Wiener process w_t and appropiately measurable $A-$valued process (a_t, b_t).

The following is a corollary of Theorem 5.4.3.

Corollary 5.4.5. *Let $\Gamma \subset \mathbb{R}^k$ be a bounded closed set and denote by $\Pi(\Gamma)$ the set of all $\pi \in \Pi$ such that*

$$\left(\mathbf{E}^\pi G_1(x_\cdot^\pi), \ldots, \mathbf{E}^\pi G_k(x_\cdot^\pi)\right) \in \Gamma. \tag{5.4.35}$$

Let $m \geqslant 0$ and assume that on $\mathcal{W}^d$ there are real-valued continuous functions $H(x.)$, $G_1(x.)$, $\ldots$, $G_k(x.)$ and a constant $K > 0$ such that for any $x. \in \mathcal{W}^d$

$$|H| + |G_1| + \ldots + |G_n| \leqslant K \left(1 + \sup_{t \in [0,T]} |x_t|\right)^m.$$

Assume that $\Pi(\Gamma) \neq \varnothing$. Then there exists $\pi_0 \in \Pi(\Gamma)$ such that

$$\mathbf{E}^{\pi_0} H(x_\cdot^{\pi_0}) = \sup_{\pi \in \Pi(\Gamma)} \mathbf{E}^\pi H(x_\cdot^{\pi_0}). \tag{5.4.36}$$

Proof. We follow [32]. Let $\{\pi_n\}_{n \geqslant 1} \subset \Pi$ be a maximising sequence of (5.4.36), by Theorem 5.4.3 there is a $\pi^\star \in \Pi$ such that $\mathbb{Q}^{\pi_{n_k}} \Rightarrow \mathbb{Q}^\star$ weakly, without loss of generality we can assume that weak convergence holds for $\{\pi_n\}$. It suffices to prove that

$$\mathbf{E}^{\pi_n} H(x_\cdot^{\pi_n}) \to \mathbf{E}^{\pi_0} H(x_\cdot^{\pi_0}) \text{ and } \mathbf{E}^{\pi_n} G_i(x_\cdot^{\pi_n}) \to \mathbf{E}^{\pi_0} G_i(x_\cdot^{\pi_0}). \tag{5.4.37}$$

If H and G_i were bounded, then the claim would follow by the definition of weak convergence. As this is not the case, first define the truncated functionals $H_r(x_\cdot^\pi) := (-r) \vee H(x_\cdot^\pi) \wedge r$ we see that

$$\lim_{r \to \infty} \sup_{\pi \in \Pi} \mathbf{E}^\pi |H_r(x_\cdot^\pi) - H(x_\cdot^\pi)| = 0. \tag{5.4.38}$$

We denote by $x_T^\star = \sup_{s\leqslant T}|x_s|$. Notice that

$$r\left|H_r\left(x_\cdot^\pi\right)-H\left(x_\cdot^\pi\right)\right| = r\Big|\left(r-H\left(x_\cdot^\pi\right)\right)\mathbb{I}_{(H(x.)>r)}\Big| \leqslant \Big(r^2+r\left|H\left(x_\cdot^\pi\right)\right|\Big)\mathbb{I}_{(H(x.)>r)}, \tag{5.4.39}$$

$$\leqslant 2\left|H\left(x_\cdot^\pi\right)\right|^2\mathbb{I}_{(H(x.)>r)} \leqslant 2K\left(1+\left(x^\pi\right)_T^\star\right)^{2m}, \tag{5.4.40}$$

then

$$\Big|H_r\left(x_\cdot^\pi\right)-H\left(x_\cdot^\pi\right)\Big| \leqslant \frac{2K}{r}\Big(1+\left(x^\pi\right)_T^\star\Big)^{2m} \text{ and } \mathbf{E}^\pi\Big|H_r\left(x_\cdot^\pi\right)-H\left(x_\cdot^\pi\right)\Big| \leqslant \frac{2K}{r}\Big(1+\mathbf{E}\left(x^\pi\right)_T^\star\Big)^{2m}. \tag{5.4.41}$$

and taking supremum over Π,

$$\sup_{\pi\in\Pi}\mathbf{E}^\pi\Big|H_r\left(x_\cdot^\pi\right)-H\left(x_\cdot^\pi\right)\Big| \leqslant \frac{2K}{r}\left(1+\sup_{\pi\in\Pi}\mathbf{E}^\pi\left[1+\left(\sup_{t\leqslant T}\left|x_t^\pi\right|\right)^{2m}\right]\right).$$

By previous moment estimates in (5.4.9) we have $\sup_{\pi\in\Pi}\mathbf{E}^\pi\left(\sup_{t\leqslant T}|x_t^\pi|\right)^{2m}<\infty$, then (5.4.38) holds.

Then, we claim we have (5.4.37), indeed

$$\mathbf{E}^{\pi_m}H\left(x_\cdot^{\pi_m}\right) \leqslant \mathbf{E}^{\pi_m}\Big|H_r\left(x_\cdot^{\pi_m}\right)-H\left(x_\cdot^{\pi_m}\right)\Big| + \mathbf{E}^{\pi_m}H_r\left(x_\cdot^{\pi_m}\right),$$

let $\varepsilon>0$, choose M large enough such that for any $r>M$

$$\sup_{\pi\in\Pi}\mathbf{E}^\pi\left|H_r\left(x_\cdot^\pi\right)-H\left(x_\cdot^\pi\right)\right| \leqslant \varepsilon, \tag{5.4.42}$$

hence $\mathbf{E}^{\pi_m}\Big(H\left(x_\cdot^{\pi_m}\right)\Big)$ is finite for any m and

$$\overline{\lim_{m\to\infty}}\,\mathbf{E}^{\pi_m}H\left(x_\cdot^{\pi_m}\right) \leqslant \varepsilon+\mathbf{E}^{\pi_0}H\left(x_\cdot^{\pi_0}\right), \tag{5.4.43}$$

as H_r is bounded, and by mononote convergence (letting $r\to\infty$). Let $\pi_0\in\Pi$ be a 'stochastic basis' $\pi_0 = \left(\Omega^{\pi_0},\mathcal{F}^{\pi_0},\{\mathcal{F}_t^{\pi_0}\},\mathbb{P}^{\pi_0},x_t^{\pi_0},w_t^{\pi_0}\right)$ and $\mathbb{Q}^n$ the weak limit of the law of $x^{\pi_0}\in\pi_0$ (Theorem 5.4.3) applying a similar estimate in π_0, let ε and $r>0$ such that (5.4.42) holds,

$$\begin{aligned}\mathbf{E}^{\pi_0}H\left(x_\cdot^{\pi_0}\right) \leqslant{}& \mathbf{E}^\star\left|H\left(x^{\pi_0}\right)-H_r\left(x^{\pi_0}\right)\right| + \\ &\left|\mathbf{E}^{\pi_0}H_r\left(x^{\pi_0}\right)-\mathbf{E}^{\pi_m}H_r\left(x_\cdot^{\pi_m}\right)\right|+\mathbf{E}^{\pi_m}H_r\left(x_\cdot^{\pi_m}\right).\end{aligned} \tag{5.4.44}$$

Letting $m\to\infty$

$$\mathbf{E}^{\pi_0}H\left(x_\cdot^{\pi_0}\right) \leqslant \mathbf{E}^{\pi_0}\left|H\left(x^{\pi_0}\right)-H_r\left(x^{\pi_0}\right)\right|+\varepsilon+\underline{\lim_{m\to\infty}}\,\mathbf{E}^{\pi_m}H\left(x_\cdot^{\pi_m}\right). \tag{5.4.45}$$

Letting now, $r\to\infty$

$$\mathbf{E}^{\pi_0}H\left(x_\cdot^{\pi_0}\right) \leqslant 2\varepsilon+\underline{\lim_{m\to\infty}}\,\mathbf{E}^{\pi_m}H\left(x_\cdot^{\pi_m}\right). \tag{5.4.46}$$

The same argument applies to G_i for every $i=1,\ldots,k$ and as Γ is a closed set, this implies $\pi_0\in\Pi(\Gamma)$. $\square$

We shall see that actually a version of Corollary 5.4.5 holds for functionals such as those describing behavioural investor preferences, at least under a type of diffusion models. This is the content of chapter 6.

Chapter 6

BEHAVIOURAL OPTIMAL INVESTMENT FOR DIFFUSION MARKET MODELS

1 Introduction

The present chapter treats preferences of cumulative prospect theory (CPT), [30, 57], where an "S-shaped" u is considered (i.e. convex up to a certain point and concave from there on) and when the time parameter is 'continuous'. Also, distorted probability measures are applied for calculating the utility of a given position with respect to a (possibly random) benchmark G. Due to the Kahneman and Tversky contributions in the understanding of investor's behaviour under uncertainty, the problem of optimal investment in CPT is a relevant subject in mathematical finance. As mentioned before, such a theory of preference has explained satisfactorily many of the paradoxes arising in one of the cornerstones of Quantitative Finance, Expected Utility Theory (EUT). Continuous-time studies have hitherto assumed a complete market model, [5, 28, 12, 10, 45]. Only very specific types of incomplete continuous-time models have been treated to date (finite mixtures of complete models; the case where the price is a martingale under the physical measure; the case where the market price of risk is deterministic), see [46, 41]. This chapter is based upon the preprint [43] and to the best of our knowledge, is a contribution to the problem of optimal investment using a weak-control approach. With this approach we are able to obtain fairly general results, we consider an incomplete market model of a diffusion type in which assets prices depend on economic factors. Our main result asserts, under some conditions, the existence of an 'optimal strategy' when the source of uncertainty of the economic factors is independent of that of the investment and the rate of return is non-negative. The independence condition is not realistic as it does not allow a leverage effect (see [7]).
We propose a further generalisation and thus, our approach also includes models where the factor may have non-zero correlation with the investment. These results open the door for further generalizations. Finally we remark that our ideas can be easily adapted to other types of preferences, such as rank-dependent utility [40] or [14] acceptability indices.

In the next section definitions and notation related to the problem of behavioural optimal investment are presented. Many of the definitions are based on the following references [32] and [11].

Unfortunately, most of the techniques developed in the literature to find optimal policies rely on either the Markovian nature of the problem or on convex duality. These are no longer applicable under behavioural criteria. For this reason, we shall consider a weak-type formulation of the control problem associated with optimal investment (section 6.2). The chapter is organised as follows. In the next section we describe the model and state our assumptions, then the problem of optimal investment under behavioural criteria is described. In section 3, we state our main result, this is based on [43]. Section 4 aims at describing the weak formulation of the problem and providing a detailed proof of the main theorem, we combine the theory developed in previous chapters. In section 5 we include the proofs of auxiliary results that were applied in section 4. Finally, in section 6 we give an extension that enables us to include correlation in the driving sources of uncertainty.

2 The setting: market and preferences

Fix a finite horizon $T > 0$. We consider a financial market consisting of a risky asset, whose discounted price $\{S_t\}_{t\in[0,T]}$ depends on economic factors. These factors are described by a $d-$dimensional stochastic processes $\{Y_t\}_{t\in[0,T]}$ solving the equation

$$dY_t = \nu_t(Y_\cdot)\,dt + \kappa_t(Y_\cdot)\,dB_t, \text{ and } Y_0 = y, \tag{6.2.1}$$

the stock price process is given by the solution of the equation

$$dS_t = \theta_t(Y_\cdot)\,S_t dt + \lambda_t(Y_\cdot)\,S_t dW_t \text{ and } S_0 = s > 0, \tag{6.2.2}$$

with B, W independent and standard Wiener processes and $W = \{W_t\}_{t\in[0,T]}$ is an $\mathbb{R}-$valued Wiener process.
We also assume that there is a riskless asset of constant price equal to 1. These shall be more specific later in this section. Stochastic volatility models provide prime examples of financial market models with dynamics (6.2.1) and (6.2.2), other potential applications are problems such as optimal investment with habit formation and assets depending on risky bond yields or other underlying whose evolution is stochastic, as well.

The investor trades in the risky and riskless assets, investing a proportion $\phi_t \in [0,1]$ of his wealth into the risky asset at time t. This leads to the following equation for the wealth of the investor at time t:

$$dX_t = \phi_t\theta_t(Y_\cdot)\,X_t dt + \phi_t\lambda_t(Y_\cdot)\,X_t dW_t \text{ and } X_0 = x, \tag{6.2.3}$$

where $x > 0$ is the investor's initial capital.

Borrowing and short selling are not allowed, hence ϕ_t is a process taking values in $[0,1]$. We note that, in this model, the risky asset's price has no influence on the economic factors. We will see in section 6.7 below how this assumption can be weakened.

We will need certain closedness results on the laws of Itô processes from [32], described in chapter 5, hence it is necessary to work in the setting of the so-called 'weak-controls', where the underlying probability space is not fixed. We shall apply the ideas of the last section.

We first set out the requirements for the coefficients in (6.2.1), (6.2.3). Let $\mathcal{W}_T^d$ denote the family of $\mathbb{R}^n$-valued continuous functions on $[0,T]$.

Denote by $x_t : \mathcal{W}_T^d \to \mathbb{R}^d$ the projections $x_t(w.) = w_t$ and define the σ-algebras $\mathscr{N}_t = \sigma(\{x_s : s \leqslant t\})$, and $\mathscr{N} = \sigma(\{x_s : s \leqslant T\})$.

Definition 6.2.1. Let $\nu(t,y.) : [0,T]\times\mathcal{W}_T^d \to \mathbb{R}^d$ be such that the restriction of ν to $[0,t]\times\mathcal{W}_t^d$ is $\mathcal{B}([0,t])\otimes\mathscr{N}_t$-measurable, for any $t\in[0,T]$. We will use the notations $\nu_t(y.)$ or $\nu(t,y.)$.

Similarly, we define the coefficients θ, λ, κ with the same measurability properties as ν, but with values in $\mathbb{R}_+$, $\mathbb{R}$ and S_+^d, respectively, where S_+^d denotes the set of real, symmetric and positive semidefinite $d\times d$ matrices.

As we are interested in weak solutions, we shall define investment strategies in a similar fashion.

Definition 6.2.2. An investment stategy π is given by the following collection:

$$\pi := \Big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, X_t, Y_t, \phi_t, (B_t, W_t), (x,y)\Big),$$

with $x > 0$ and $y\in\mathbb{R}^d$, where

(a) $\Big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\Big)$ is a complete filtered probability space whose filtration satisfies the usual conditions;

(b) the process $(W_t, B_t)_{t\in[0,T]}$ is a standard $d+1$-dimensional $\mathcal{F}_t$-Wiener process;

(c) $\phi_t : \Omega\times[0,T]\to[0,1]$ is $\mathcal{F}\otimes\mathcal{B}([0,T])$-measurable and $\mathcal{F}_t$-adapted;

(d) on the filtered probability space X_t, Y_t are $\mathcal{F}\otimes\mathcal{B}([0,T])$-measurable and $\mathcal{F}_t$-adapted processes such that

$$Y_t = y + \int_0^t \nu_s(Y.)\,ds + \int_0^t \kappa_s(Y.)\,dB_s, \tag{6.2.4}$$

$$X_t = x + \int_0^t \phi_s\theta_s(Y.)\,X_s ds + \int_0^t \phi_s\lambda_s(Y.)\,X_s dW_s, \tag{6.2.5}$$

for $0\leqslant t\leqslant T$.

In other words, $\Big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, X_t, Y_t, (B_t, W_t), (x,y)\Big)$ is a weak solution of the system of equations (6.2.1), (6.2.3) with initial condition (x,y). The process ϕ_t represents a ratio of investment in the risky asset, it is measurable and $\mathcal{F}_t-$adapted. We do not consider the price process S_t from (6.2.2) at all since it is enough to work with the 'controlled dynamics' X_t.

When needed, we will use the notation X^π, Y^π, etc. to indicate that the object we mean belongs to π. We denote the family of all investment strategies in the sense of definition 6.2.2 by Π.

Assumption 6.2.1. *The functional θ is non-negative i.e. $\theta(t,y.)\geqslant 0$ for all $t\in[0,T]$ and $y\in\mathcal{W}_T^d$.*

Remark 6.2.3. Assumption 6.2.1 means simply that the return of the risky asset must be non-negative. This looks rather a harmless assumption. On the other hand, as mentioned before, (d) in Definition 6.2.2 is stringent. It excludes the 'leverage effect' where the volatility and the stock prices have (negative) correlation. This condition can be relaxed, see section 6.7.

We now present the framework of optimal investment under CPT, as proposed in [57]. We follow [41] and [11].

The investor assesses strategies by means of utilities on gains and losses, which are described in terms of functions $u_{\pm} : \mathbb{R}_+ \to \mathbb{R}$, by a reference point G and functions $w_{\pm} : [0,1] \to [0,1]$. The latter functions $w_{\pm}$ are included with the aim of explaining the distortions of her perception on the "likelihood" of her gains and losses.

We model the reference point G by a real-valued random variable G, as explained before in Section 3 the investor uses the benchmark G to asses the portfolio outcomes.
The quantity G depends on economic factors as follows: let us denote by F a fixed deterministic functional $F : \mathcal{W}_T^d \to \mathbb{R}_+$ which is $\mathscr{N}_T-$measurable. As the probability space is not fixed, for each $\pi \in \Pi$, we define the corresponding reference point by $G^\pi := F\left(Y_\cdot^\pi\right)$. That is, we assume that the benchmark is a non-negative functional of the economic factors.

For any strategy $\pi \in \Pi$, we define the functionals

$$V_+(\pi) := \int_0^\infty w_+\left(\mathbb{P}^\pi\left(u_+\left(\left(X_T^\pi - G^\pi\right)_+\right) > t\right)\right) dt, \tag{6.2.6}$$

and

$$V_-(\pi) := \int_0^\infty w_-\left(\mathbb{P}^\pi\left(u_-\left(\left(X_T^\pi - G^\pi\right)_-\right) > t\right)\right) dt. \tag{6.2.7}$$

The optimal portfolio problem for an investor under CPT consists in maximising the following performance functional:

$$V(\pi) := V_+(\pi) - V_-(\pi), \tag{6.2.8}$$

which is defined provided that at least one of the summands is finite. Set $\Pi' := \{\pi \in \Pi : V_-(\pi)\} < \infty$ and define

$$V := \sup_{\pi \in \Pi'} V(\pi). \tag{6.2.9}$$

The value V represents the maximal satisfaction achievable by investing in the stock and riskless asset in a CPT framework. Our purpose is to prove the existence of $\pi' \in \Pi$ such that $V(\pi') = V$.

3 Main result

We make the following assumptions. Recall the notation $y_t^\star = \sup_{s \leqslant t} |y_s|$.

Assumption 6.3.1. *The functionals κ, λ, θ and ν are uniformly bounded. Furthermore, for fixed $t \geqslant 0$ and functions $y^n, z \in \mathcal{W}_T^d$ such that $(y^n - z)_t^\star \to 0$ we have $\kappa_t(y_\cdot^n) \to \kappa_t(z_\cdot)$ and the same holds for the functionals λ, θ and ν. We mean this when we write later that "the coefficients are continuous".*

Assumption 6.3.2. *A (weak) solution of equation (6.2.4) exists and it is unique in law.*

Assumption 6.3.3. *We assume that $u_{\pm} : \mathbb{R}_+ \to \mathbb{R}_+$ and $w_{\pm} : [0,1] \to [0,1]$ are continuous,* *non-decreasing functions with* $u_\pm(0) = 0$, $w_\pm(0) = 0$, $w_\pm(1) = 1$, *and*

$$u_+(x) \leqslant k_+(x^\alpha + 1), \textit{for all } x \in \mathbb{R}_+ \tag{6.3.1}$$

$$w_+(p) \leqslant g_+ p^\gamma, \textit{ for all } p \in [0,1] \tag{6.3.2}$$

with γ, $\alpha > 0$, $k_+, g_+ > 0$ *fixed constants.*

Denote by $L^p(\Omega, \mathbb{P})$ the usual space of p-integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Assumption 6.3.4. *There is* $\vartheta > 0$ *such that* $\vartheta\gamma > 1$ *and* $G^\pi \in L^{\vartheta\gamma}(\Omega, \mathbb{P}^\pi)$ *for all* $\pi \in \Pi$.

Note that, under Assumption 6.3.2, the law of G^π is independent of π and hence Assumption 6.3.4 holds iff $G^\pi \in L^{\vartheta\gamma}(\Omega, \mathbb{P}^\pi)$ for one particular π.

In order to ensure that the functional V and the optimisation problem in (6.2.9) are defined over a non-empty set, we introduce the following assumption on u_-, the distortion function w_- and the reference point G^π.

Assumption 6.3.5. *The functions* w_-, u_- *are such that, for all* $\pi \in \Pi$,

$$\int_0^\infty w_-(\mathbb{P}^\pi(u_-(G^\pi) > y))\, dy < \infty. \tag{6.3.3}$$

Assumption 6.3.5 ensures that the set Π' is not empty. Indeed, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space where (6.2.1) has a solution Y_t. Then setting $\phi_t := 0$ and $X_t := x$ for all t,

$$\Big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, x, Y_t, 0, (B_t, W_t), (x, y)\Big),$$

belongs to Π' for each $x > 0$. Our main result can now be stated. The main theorem of this chapter is the following.

Main Theorem 6.3.1. *Under Assumptions 6.2.1, 6.3.1, 6.3.2, 6.3.3, 6.3.4 and 6.3.5 the problem* (6.2.9) *is well-posed, i.e.* $V < \infty$. *Moreover, there exists an optimal strategy* $\hat{\pi} \in \Pi$ *(see definition 6.2.2) attaining the supremum in (6.2.9), i.e.* $V = V(\hat{\pi})$.

Remark 6.3.2. Theorem 6.3.1 holds even if κ and ν are not bounded but satisfy a linear growth condition. On the other hand, continuity in Assumption 6.3.1 is essential.

4 A relaxation of the problem

We introduce a relaxation of the problem by extending the class of investment strategies given in Definition 6.2.2, we shall call this extension the class of auxiliary controls. This relaxation is introduced in order to ensure the closedness of the set of laws of the processes $(Y_\cdot, X_\cdot)$.

We follow the martingale problem formulation, thus we refer to (a_t, b_t) as the characteristics of the diffusion (Y_t, X_t), such processes are the ones involved in the martingale problem formulation of equations (6.2.4) and (6.2.5). For this, the characteristics must take values in a family of convex subsets of $S_+^{d+1} \times \mathbb{R}^{d+1}$ hence we shall consider a 'convex extension' of the set on which the characteristics (and hence of the coefficients) in equations (6.2.4) and (6.2.5) take values.

Definition 6.4.1. Denote $\mathbb{A} = S_+^{d+1} \times \mathbb{R}^{d+1}$. For any pair of continuous functions $(x_\cdot, y_\cdot) \in C([0,T]; \mathbb{R} \times \mathbb{R}^d)$ and for any $t \in [0,T]$ we define

$$A_t(x_\cdot, y_\cdot) = \left\{ (a,b) \in \mathbb{A} \middle| (a,b) = \left(\begin{pmatrix} \frac{1}{2}\kappa\kappa^\star(t,y_\cdot) & 0 \\ 0 & \frac{1}{2}m\lambda^2(t,y_\cdot)x_t \end{pmatrix}, \begin{pmatrix} \nu(t,y_\cdot) \\ l\theta(t,y_\cdot)x_t \end{pmatrix} \right), \begin{matrix} 0 \leqslant m \leqslant 1, \\ 0 \leqslant l \leqslant \sqrt{m} \end{matrix} \right\}. \tag{6.4.1}$$

Remark 6.4.2. Notice that, for any investment strategy π as in Definition 6.2.2, if $\sigma_t = \begin{pmatrix} \kappa(t,y_\cdot) & 0 \\ 0 & \phi_t\lambda(t,y_\cdot)x_t \end{pmatrix}$ and $b_t = \begin{pmatrix} \nu(t,y_\cdot) \\ \phi_t\theta(t,y_\cdot)x_t \end{pmatrix}$ then, defining $a_t = \frac{1}{2}\sigma_t\sigma_t^\star$, the pair (a_t, b_t) belongs to $A_t(x_\cdot, y_\cdot)$.

We now describe the familiy of auxiliary controls used throughout this work. It stresses the fact of having Itô processes whose characteristics belong to the convex sets $A_t(x_\cdot, y_\cdot)$ in 'a measurable way' as $t, x_\cdot$ and $y_\cdot$ vary.

Definition 6.4.3. We define a family of auxiliary controls $\overline{\Pi}$. Namely, an auxiliary control $\overline{\pi} \in \overline{\Pi}$ consists of the collection

$$\overline{\pi} := \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, X_t, Y_t, \xi_t, (x,y)\right),$$

where

(a) $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ is a complete filtered probability space whose filtration satisfies the usual conditions;

(b) $\xi_t := (B_t, W_t)$ is an $\mathbb{R}^{d+1}$-valued standard $\mathcal{F}_t$-Wiener process and W_t is an $\mathbb{R}-$valued Wiener process;

(c) there exists an $\mathbb{A}$-valued, $\mathcal{F} \otimes \mathcal{B}([0,T])$-measurable and $\mathcal{F}_t$-adapted process, denoted by (a_t, b_t), such that (d) and (e) below hold;

(d) The processes X_t and Y_t are $\mathcal{F} \otimes \mathcal{B}([0,T])$-measurable and $\mathcal{F}_t$-adapted such that a.s. for all $t \geqslant 0$;
$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} + \int_0^t \sqrt{2a_s}\,d\xi_s + \int_0^t b_s ds; \tag{6.4.2}$$

(e) for almost all $(\omega, t) \in \Omega \times [0,T]$, we have $(a_t, b_t) \in A_t(X_\cdot, Y_\cdot)$.

We will often write $X^{\overline{\pi}}$, $Y^{\overline{\pi}}$ to indicate that we mean X, Y belonging to $\overline{\pi}$.

For each $\overline{\pi} \in \overline{\Pi}$, we can define $V_\pm(\overline{\pi})$ as before and we can set $V(\overline{\pi}) := V_+(\overline{\pi}) - V_-(\overline{\pi})$ for $\overline{\pi} \in \overline{\Pi}' := \{\pi \in \overline{\Pi} : V_-(\pi) < \infty\}$.

Remark 6.4.4. The relationship between the processes a_t and b_t in $A_t(x_\cdot, y_\cdot)$ and the real-valued processes l_t and m_t with $0 \leqslant m_t \leqslant 1, 0 \leqslant l_t \leqslant \sqrt{m_t}$ is clear from Definition 6.4.3. Condition (c) in Definition 6.4.3, i.e. the measurability properties of the $\mathbb{A}$-valued process (a_t, b_t), yields the same properties for the processes l_t and m_t: (d) implies that l_t, m_t can be chosen $\mathcal{F} \otimes \mathcal{B}([0,T])$ measurable and $\mathcal{F}_t-$adapted.

Equation (6.4.2) can be rewritten as the set of equations below. Denote

$$\sigma_t = \begin{pmatrix} \kappa(t,Y_\cdot) & 0 \\ 0 & \sqrt{m_t}\lambda(t,Y_\cdot)X_t \end{pmatrix},$$

and

$$b_t = \begin{pmatrix} \nu(t, Y.) \\ l_t\theta(t, Y.) X_t \end{pmatrix}.$$

Setting $a_t := \frac{1}{2}\sigma_t\sigma_t^{\star}$,

$$Y_t = y + \int_0^t \nu_s(Y.)\, ds + \int_0^t \kappa_s(Y.)\, dB_s, \tag{6.4.3}$$

$$X_t = x + \int_0^t l_s\theta_s(Y.) X_s ds + \int_0^t \sqrt{m_s}\lambda_s(Y.) X_s dW_s. \tag{6.4.4}$$

Definition 6.4.5. Let $\overline{\pi} \in \overline{\Pi}$ be a relaxed control, we say that $\{X_t^{\overline{\pi}}\}_{t\in[0,T]}$ is a *portfolio value* process if $l_t = \sqrt{m_t}$, i.e.

$$dX_t = \sqrt{m_t}\theta(t, Y.) X_t dt + \sqrt{m_t}\lambda(t, Y.) X_t dW_t. \tag{6.4.5}$$

Remark 6.4.6. If $X_t^{\overline{\pi}}$ is a portfolio value process then taking $\phi_t = \sqrt{m_t}$ we see that

$$\left(\Omega^{\overline{\pi}}, \mathcal{F}^{\overline{\pi}}, \{\mathcal{F}_t^{\overline{\pi}}\}_{t\in[0,T]}, \mathbb{P}^{\overline{\pi}}, X_t^{\overline{\pi}}, Y_t^{\overline{\pi}}, \phi_t^{\overline{\pi}}, \left(B_t^{\overline{\pi}}, W_t^{\overline{\pi}}\right), (x, y)\right),$$

belongs to Π.

Remark 6.4.7. Suppose that we are given a $\overline{\pi} \in \overline{\Pi}$ i.e. there is a standard $d+1$-dimensional Wiener process $(B_t, W_t)_{t\geqslant 0}$ on $\left(\Omega^{\overline{\pi}}, \mathcal{F}^{\overline{\pi}}, \{\mathcal{F}_t^{\overline{\pi}}\}_{t\in[0,T]}, \mathbb{P}^{\overline{\pi}}\right)$ and processes $X_t^{\overline{\pi}}, Y_t^{\overline{\pi}}, a_t^{\overline{\pi}}, b_t^{\overline{\pi}}$ such that equations (6.4.3) and (6.4.4) hold.
Define the continuous semimartingale $M_t^{\overline{\pi}} := \int_0^t \sqrt{m_s^{\overline{\pi}}}\lambda_s(Y.)\, dW_s + \int_0^t l_s^{\overline{\pi}}\theta_s(Y.)\, ds$. Then we can rewrite equation (6.4.3) as

$$X_t = x + \int_0^t X_s dM_s^{\overline{\pi}}. \tag{6.4.6}$$

Equation (6.4.6) has a unique strong solution given by the stochastic exponential, and then $X_t^{\overline{\pi}}$ is given by

$$X_t^{\overline{\pi}} = x \exp\left\{\int_0^t \sqrt{m_s}\lambda_s\left(Y_\cdot^{\overline{\pi}}\right) dW_s^{\overline{\pi}} + \int_0^t \left[l_s\theta_s\left(Y_\cdot^{\overline{\pi}}\right) - \frac{1}{2}m_s\lambda_s^2\left(Y_\cdot^{\overline{\pi}}\right)\right] ds\right\}, \tag{6.4.7}$$

and this process is positive $\mathbb{P}^{\overline{\pi}}$-a.s.

5 Compactness of laws and related results

Lemma 6.5.1. *Let $M = \max\{\|\kappa\|_\infty, \|\lambda\|_\infty, \|\theta\|_\infty, \|\nu\|_\infty\}$ then the set $A_t(x., y.)$ is convex, closed and bounded, where the bound depends on M and x_t only.*

Proof. Notation $|\cdot|$ will refer to Euclidean norms of varying dimensions. For simplicity, we assume $d = 1$. Notice that

$$|(\sigma_t, b_t)| = \left(\|\kappa_t\|^2(y.) + m^2 \cdot \lambda_t^2(y.) \cdot x_t^2 + |\nu_t|^2(y.) + l^2\theta_t^2(y.)\, x_t^2\right)^{1/2},$$

hence

$$|(\sigma_t, b_t)| \leqslant M'(1 + |x_t|). \tag{6.5.1}$$

It is clear that the set is closed. For a fixed t, $x.$ and $y.$ the set is bounded. Indeed, let

$(a_t, b_t) \in A_t(x., y.)$ and

$$(a_t, b_t) = \begin{pmatrix} \frac{1}{2}\kappa^2(t, y.) & 0 \\ 0 & \frac{1}{2}m\lambda^2(t, y.) \cdot x_t^2 \end{pmatrix}, \begin{pmatrix} \nu(t, y.) \\ l\theta(t, y.)\, x_t \end{pmatrix},$$

we have

$$|(a_t, b_t)| = \left(\frac{1}{4} \cdot \left(\kappa_t^2(y.)\right)^2 + \frac{1}{4}\left(m \cdot \lambda_t^2(y.) \cdot x_t^2\right)^2 + \nu_t^2(y.) + l^2\theta_t^2(y.)\, x_t^2\right)^{1/2},$$

so

$$|(a_t, b_t)| \leqslant \left(\frac{1}{4}M^4 + \frac{1}{4}M^4 \cdot \left(x_t^2\right)^2 + M^2 + M^2 x_t^2\right)^{1/2},$$

which leads to $|(a_t, b_t)| \leqslant \frac{1}{2}(M+1)^2 + M^2 x_t^2$.

In particular,

$$\|A_t(x., y.)\| := \max\{|(a_t, b_t)| : (a_t, b_t) \in A_t(x., y.)\} \leqslant K\left(1 + |x_t|^2\right), \tag{6.5.2}$$

for some $K \geqslant 0$.

The set $A_t(x., y.)$ is also convex. Indeed, let $(\alpha, b), (\gamma, c) \in A_t(x., y.)$ then, for $0 \leqslant \mu \leqslant 1$,

$$\mu(\alpha, b) + (1-\mu)(\gamma, c) = \begin{pmatrix} \frac{1}{2}\kappa_t^2(y.) & 0 \\ 0 & \frac{1}{2}(\mu m + (1-\mu)n)\,\lambda_t^2(y.)\, x_t^2 \end{pmatrix}, \begin{pmatrix} \nu_t(y.) \\ (\mu l + (1-\mu)p)\,\theta_t(y.)\, x_t \end{pmatrix},$$

with $0 \leqslant m, n \leqslant 1$, $0 \leqslant l \leqslant \sqrt{m}$ and $0 \leqslant p \leqslant \sqrt{n}$. Clearly, $\mu l + (1-\mu)p \leqslant \sqrt{\mu m + (1-\mu)\, n}$, by concavity of $\sqrt{\cdot}$. □

Notice that the last estimates hold true if a linear growth condition is satisfied by κ and ν, the bound would depend on $|x_t|$ and $|y_t|$.
In order to deal with (semi)continuity issues related to the family of sets defined in Definitions 6.4.1 and (6.4.1), the support function of the familiy of sets $A_t(x., y.)$ is now considered.
We denote for all $u \in \mathbb{R}^{(d+1)(d+1)}$, $v \in \mathbb{R}^{d+1}$ and $t \in [0, T]$,

$$F_t(x., y.)(u, v) = \max\left\{\sum_{i,j=1}^{d+1} a_{ij}u_{ij} + \sum_{j=1}^{d+1} b_j v_j \;:\; (a, b) \in A_t(x., y.)\right\}. \tag{6.5.3}$$

Under Assumption 6.3.1, for fixed $t \geqslant 0$ and (u, v), the support function $(x., y.) \to F_t(y., x.)(u, v)$ is continuous.
In particular, the set $A_t(x., y.)$ is upper-semicontinuous in the sense of Assumption 3.1 iii) in [32]. It is also clear that, for fixed $u, v \in \mathbb{A}$, $F_t(u, v, \zeta.)$ is a Borel function on $[0, T] \times C\left([0, T]; \mathbb{R}^{d+1}\right)$.

We first present some moment estimates which will, in particular, guarantee tightness for the family of the laws of $(X^{\overline{\pi}}, Y^{\overline{\pi}})$, $\overline{\pi} \in \overline{\Pi}$ in $C\left([0, T]; \mathbb{R}^{d+1}\right)$.

Proposition 6.5.2. *For the ease of reference we denote $\zeta_t = (Y_t, X_t)$. For any $m > 0$,*

$$\sup_{\overline{\pi} \in \overline{\Pi}} \mathbf{E}^{\overline{\pi}}\left[\sup_{t \in [0,T]} |\zeta_t^{\overline{\pi}}|^m\right] < \infty. \tag{6.5.4}$$

Proposition 6.5.3. *Under Assumption 6.3.1, let $\overline{\pi} \in \overline{\Pi}$ and $\left(Y_t^{\overline{\pi}}, X_t^{\overline{\pi}}\right)$ their associated processes solving (6.4.3) and (6.4.4). Then, there exist a constant $K_T > 0$ not depending on $\overline{\pi} \in \overline{\Pi}$, such that for any $\eta > 0$ and $s, t \in [0, T]$,*

$$\mathbf{E}^{\overline{\pi}} \|\zeta_t - \zeta_s\|^\eta \leqslant K_T \, |t - s|^{\frac{\eta}{2}}. \tag{6.5.5}$$

See section 6.8.1 for a standard proof of both propositions above. A well-known result on tightness of measures on $C\left([0,T]; \mathbb{R}^{d+1}\right)$ gives the following corollary.

Corollary 6.5.4. *Let Assumption 6.3.1 be in force. Let $\{\overline{\pi}_n\} \subset \overline{\Pi}$. The set of laws of the process $\zeta_\cdot^{\overline{\pi}_n}$ on $C\left([0,T]; \mathbb{R}^{d+1}\right)$ is relatively weakly compact.*

Remark 6.5.5. Alternatively, if π_n is a sequence of auxiliary controls, we can easily check the following estimates

$$\langle b_t^n(\zeta_\cdot), \zeta_t \rangle + tr\left(a_t^n(\zeta_\cdot)\right) \leqslant 2M\left(1 + |\zeta_t|^2\right), \tag{6.5.6}$$

and

$$|b_t^n(\zeta_\cdot)| \leqslant \left(|\nu_t(Y_\cdot)|^2 + l_t^2 \theta_t^2(Y_\cdot) X_t^2\right)^{\frac{1}{2}} \leqslant M + M|X_t| \leqslant 2M\left(1 + |\zeta_t|^2\right), \tag{6.5.7}$$

$$tr\left(a_t^n(\zeta_\cdot)\right) \leqslant \frac{1}{2}\left[|\kappa_t(Y_\cdot)|^2 + m_t \lambda_t^2(Y_\cdot) X_t^2\right] \leqslant \frac{1}{2} M^2 \left[1 + |\zeta_t|^2\right]. \tag{6.5.8}$$

Thus,

$$|b_t^n(\zeta_\cdot)| + tr\left(a_t^n(\zeta_\cdot)\right) \leqslant K\left(1 + |\zeta_t|^2\right), \tag{6.5.9}$$

with $K = \frac{1}{2}M^2 + 2M$. Let us define $L(r,t) := K\left(1 + r^2\right)$, it is an increasing function of $r \in [0, \infty)$ thus, Assumption 4.2.2 holds with this choice. On the other hand, (6.5.6) ensures that Assumption 4.2.3 holds; indeed, (6.5.6) implies that $L_n(t)$ in Assumption 4.2.3 is deterministic and (4.2.63) holds trivially. These estimates allow to prove that the laws of ζ_t are tight by a direct application of Theorem 4.2.38.

Now we restate Theorem 3.2 of [32], Theorem 5.4.3 in our setting, which will provide weak compactness of the laws of $\left(x^{\overline{\pi}}, y^{\overline{\pi}}\right)$.

Theorem 6.5.6. *Let Assumption 6.3.1 be in force. Denote by $\mathbb{Q}^{\overline{\pi}}$ the distribution of $\zeta_\cdot^{\overline{\pi}}$ on $C\left([0,T]; \mathbb{R}^{d+1}\right)$. Then the set $\left\{\mathbb{Q}^{\overline{\pi}} : \overline{\pi} \in \overline{\Pi}\right\}$ is sequentially weakly compact.*

Proof. It follows from the above discussions that Assumption 5.4.1 ii) and iii) (Assumption 3.1 in [32]) hold in the present case. One does not have Assumption 5.4.1 i), i.e. Assumption 3.1. in 5.4.1 though (linear growth condition on $\|A_t(X_\cdot, Y_\cdot)\|$), there is a quadratic growth instead, see (6.5.2). But, as Corollary 6.5.4 or Remark 6.5.5 show, this is still sufficient to get tightness (and hence relative weak compactness by Prokhorov theorem) of the sequence $\mathbb{Q}^{\overline{\pi}_n}$ in our setting. Then one can check that the proof of Theorem 3.2 in [32], Theorem 5.4.3 goes through and we can conclude. ☐

The next lemma shows that, to any auxiliary control $\overline{\pi}$ in the sense of Definition 6.4.3, we can associate an *investment strategy* (in the sense of Definition 6.2.2) with higher value function.

Lemma 6.5.7. *Let*

$$\overline{\pi} = \left(\Omega^{\overline{\pi}}, \mathcal{F}^{\overline{\pi}}, \{\mathcal{F}_t^{\overline{\pi}}\}_{t \in [0,T]}, \mathbb{P}^{\overline{\pi}}, \left(X_t^{\overline{\pi}}, Y_t^{\overline{\pi}}\right), \left(B_t^{\overline{\pi}}, W_t^{\overline{\pi}}\right), (x,y)\right) \in \overline{\Pi}.$$

Then a solution to

$$dY_t = \nu_t (Y_\cdot)\, dt + \kappa_t (Y_\cdot)\, dB_t, \quad Y_0 = y, \tag{6.5.10}$$

$$d\hat{X}_t = \sqrt{m_t}\theta_t (Y_\cdot)\, \hat{X}_t dt + \sqrt{m_t}\lambda_t (Y_\cdot)\, \hat{X}_t dW_t, \quad X_0 = x, \tag{6.5.11}$$

exists on the same filtered probability space and $\hat{X}_T \geqslant X_T^{\overline{\pi}}$ a.s. Furthermore, $\hat{X}_t$ is a portfolio value process.

Proof. Let us define

$$Z_t := \exp\Big(- \int_0^t \Big(l_s^{\overline{\pi}} - \sqrt{m_s^{\overline{\pi}}} \Big) \Big(X_\cdot^{\overline{\pi}}, Y_\cdot^{\overline{\pi}} \Big) \theta_s \left(Y_\cdot^{\overline{\pi}} \right) ds \Big)$$

and set $\hat{X}_t := Z_t X_t^{\overline{\pi}}$. Itô's formula shows that $\hat{X}_t$ indeed verifies (6.5.11). Since $\theta_t \geqslant 0$ was assumed, we get that $Z_t \geqslant 1$ hence $\hat{X}_t \geqslant X_t^{\overline{\pi}}$, for all t. □

6 Proof of Theorem 6.3.1

Proof. Let $y > 0$. By (6.3.2) and (6.3.1),

$$w_+ \left(\mathbb{Q}^\pi \left(u_+ \left((X_T^\pi - G^\pi)_+ \right) > y \right) \right) \leqslant g_+ \left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^\alpha > \frac{y}{k_+} - 1 \right) \right]^\gamma .$$

Hence,

$$\begin{aligned} V_+ (\pi) &\leqslant g_+ \int_0^\infty \left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^\alpha > \frac{y}{k_+} - 1 \right) \right]^\gamma dy = \\ &= g_+ \left(1 + \int_{k_+}^\infty \left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^\alpha > \frac{y}{k_+} - 1 \right) \right]^\gamma dy \right), \end{aligned}$$

$$\int_{k_+}^\infty \left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^\alpha > \frac{y}{k_+} - 1 \right) \right]^\gamma dy \leqslant k_+ \int_0^\infty \left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^\alpha > x \right) \right]^\gamma dx. \tag{6.6.1}$$

If $x \geqslant 1$, applying Chebyshev's inequality and Assumption 6.3.4,

$$\left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^\alpha > x \right) \right]^\gamma = \left[\mathbb{Q}^\pi \left((X_T^\pi - G^\pi)_+^{\alpha\vartheta} > x^\vartheta \right) \right]^\gamma \leqslant \frac{\left[\mathbf{E}^\pi \left(X_T^\pi - G^\pi \right)_+^{\alpha\vartheta} \right]^\gamma}{x^{\vartheta\gamma}} \leqslant M^\gamma \frac{1}{x^{\vartheta\gamma}}, \tag{6.6.2}$$

where $M = \sup_\pi \mathbf{E}^\pi \left(X_T^\pi \right)_+^{\alpha\vartheta} < \infty$ (note that $G \geqslant 0$), by Proposition 6.5.2. Note that $1/x^{\vartheta\gamma}$ is integrable on $[1, \infty)$.

Hence the problem is well-posed since $V(\pi) \leqslant V_+(\pi)$ for all $\pi \in \Pi'$ and we have just seen that the latter has an upper bound independent of π.

By Theorem 6.5.6 the set of laws $\{\mathbb{Q}^\pi\}$, $\pi \in \overline{\Pi}$ of the processes $\zeta^\pi = (X_\cdot^\pi, Y_\cdot^\pi)$ is relatively compact in the weak topology. Let $\{\pi^n\}$ be sequence of weak controls $\pi^n \in \overline{\Pi}'$ such that

$$V(\pi^n) \to \sup_{\pi \in \overline{\Pi}'} V(\pi), \quad n \to \infty. \tag{6.6.3}$$

There is a subsequence of $\{\pi^n\}$ denoted by $\left\{\pi^k\right\}$ such that $\mathbb{Q}^{\pi_k} \Rightarrow \mathbb{Q}^{\pi^\star}$ as $k \to \infty$ and

$\pi^\star \in \overline{\Pi}$.

By Skorokhod's theorem there is a probability space, that will be denoted by $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}\right)$ and random variables $\tilde{X}^k_\cdot, \tilde{Y}^k_\cdot : \left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}\right) \to \mathcal{W}^1_T$, $\mathcal{W}^d_T$, respectively, such that the law of $\left(\tilde{X}^k_\cdot, \tilde{Y}^k_\cdot\right)$ equals $\mathbb{Q}^{\pi_k}$ and $\tilde{X}, \tilde{Y} : \left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}\right) \to \mathcal{W}^1_T$, $\mathcal{W}^d_T$ with law equal to $\mathbb{Q}^{\pi^\star}$ such that $\tilde{X}^k \to \tilde{X}$, $\tilde{Y}^k \to \tilde{Y}$ a.s. in the uniform norm.

By Assumption 6.3.2, $\tilde{Y}^k$ and $\tilde{Y}$ have the same law and $\tilde{Y}^k \to \tilde{Y}$ in probability (even a.s.). By Theorem 2.2.6 or Théorème 1 in [3] $F\left(\tilde{Y}^k\right) \to F\left(\tilde{Y}\right)$ in probability.

By continuity of $u_\pm$ and the projection $p_T(f) := f(T)$, $f \in \mathcal{W}^1_T$, then $u_\pm\left(\left(\tilde{X}^k_T - F\left(\tilde{Y}^k\right)\right)_\pm\right) \to u_\pm\left(\left(\tilde{X}_T - F\left(\tilde{Y}\right)\right)_\pm\right)$ in probability.

It follows that, denoting by D the set of discontinuity points of the cumulative distribution functions of $u_\pm\left(\left(\tilde{X}_T - F\left(\tilde{Y}\right)\right)_\pm\right)$, for any $y \in \mathbb{R}\backslash D$ we have

$$\mathbb{Q}^{\pi_k}\left(u_\pm\left((X^{\pi_k}_T - G^{\pi_k})_\pm\right) > y\right) \to \mathbb{Q}^{\pi^\star}\left(u_\pm\left(\left(X^{\pi^\star}_T - G^{\pi^\star}\right)_\pm\right) > y\right).$$

as $k \to \infty$.

Since $w_\pm$ are continuous, also

$$w_\pm\left(\mathbb{Q}^{\pi_k}\left(u_\pm\left((X^{\pi_k}_T - G^{\pi_k})_\pm\right) > y\right)\right) \to w_\pm\left(\mathbb{Q}^{\pi^\star}\left(u_\pm\left(\left(X^{\pi^\star}_T - G^{\pi^\star}\right)_\pm\right) > y\right)\right),$$

for $y \notin D$. By Fatou's lemma, and by (6.6.2) (just as it was proved in chapter 3 Theorem 3.3.16).

$$\int_0^\infty w_-\left(\mathbb{Q}^{\pi^\star}\left(u_-\left(\left(X^{\pi^\star}_T - G^{\pi^\star}\right)_-\right) > y\right)\right) dy = \lim_{k\to\infty} \int_0^\infty w_-\left(\mathbb{Q}^{\pi_k}\left(u_-\left((X^{\pi_k}_T - G^{\pi_k})_-\right) > y\right)\right) dy.$$

It follows that $V(\pi^\star) = \sup_{\pi\in\overline{\Pi}'} V(\pi)$. It is also clear that $\pi^\star \in \overline{\Pi}'$. Let (a_t, b_t) be the $\mathbb{A}$-valued processes associated to $\pi^\star \in \overline{\Pi}$ as in Definition 6.4.3. By Lemma 6.5.7 there is

$$\pi' = \left(\Omega^{\pi^\star}, \mathcal{F}^{\pi^\star}, (\mathcal{F}^{\pi^\star}_t)_{t\geqslant 0}, \mathbb{P}^{\pi^\star}, X^{\pi'}, Y^{\pi^\star}, \left(B^{\pi^\star}, W^{\pi^\star}\right), (x, y)\right)$$

which is a portfolio value process in the sense of Definition 6.4.5 and for which

$$u_+\left(\left(X^{\pi'}_T - G^{\pi'}\right)_+\right) \geqslant u_+\left(\left(X^{\pi^\star}_T - G^{\pi^\star}\right)_+\right),$$

notice that $G^{\pi'} = G^{\pi^\star}$ and $u_-\left(\left(X^{\pi^\star}_T - G^{\pi^\star}\right)_-\right) \geqslant u_-\left(\left(X^{\pi'}_T - G^{\pi'}\right)_-\right)$ also, so $V(\pi') \geqslant \sup_{\pi\in\overline{\Pi}'} V(\pi)$. Thus, recalling Remark 6.4.6, the investment strategy

$$\hat{\pi} = \left(\Omega^{\pi'}, \mathcal{F}^{\pi'}, \mathbb{P}^{\pi'}, \left\{\mathcal{F}^{\pi'}_t\right\}_{t\in[0,T]}, X^{\pi'}_t, Y^{\pi'}_t, \sqrt{m^{\pi'}_t}, (B^{\pi'}_t, W^{\pi'}_t), (x, y)\right)$$

is optimal i.e.

$$\sup_{\pi\in\Pi'} V(\pi) \leqslant V(\pi^*) \leqslant V(\hat{\pi}), \tag{6.6.4}$$

and obviously, $\hat{\pi} \in \Pi'$. □

7 Extensions

8 Description of the market model

In this section, we extend the market model that was developed in the last sections, by allowing the portfolio value process to influence the factor process Y_t, the influence in our model is 'additive'. Furthermore, a riskless asset with deterministic interest rate r_t at time t is included. For simplicity we will assume that the factor process Y takes values in $\mathbb{R}$. The computations are easily extended to the case when the factor process $\{Y_t\}_{t\in[0,T]}$ is multi-dimensional.

Definition 6.8.1. Let $\nu(t, y.)$ be an $\mathbb{R}$-valued process, such that the restriction of ν to $[0,t] \times \mathcal{W}_t^1$ is $\mathcal{B}([0,t]) \otimes \mathcal{N}_t-$measurable, for any $t \in [0,T]$. We will use both the notations $\nu_t(y.)$ and $\nu(t, y.)$.

Similarly, we define the coefficients $\theta, \lambda, \rho, \kappa$ to be $\mathcal{B}([0,T]) \otimes \mathcal{N}_T-$measurable and $\mathcal{N}_t-$adapted with values in $\mathbb{R}$.

In this case, the stochastic differential equations of the optimal investment model are given by

$$dY_t = \nu(t, Y.)\, dt + \kappa(t, Y.)\, dB_t + \rho(t, X.)\, dX_t, \quad Y_0 = y, \tag{6.8.1}$$

$$dX_t = \phi_t \theta(t, Y.)\, X_t dt + \phi_t \lambda(t, Y.)\, X_t dW_t + (1 - \phi_t)\, r_t X_t dt, \quad X_0 = x. \tag{6.8.2}$$

where $\phi_t \in [0,1]$ represents the proportion of wealth invested in the stock, Y is an economic factor and X is the wealth given by the portfolio strategy ϕ. The set Π is defined analogously to Definition 6.2.2.

Assumption 6.8.1. *For all $t \geqslant 0$, the growth rate of the stock is greater than the growth rate of the bond, i.e.*

$$\theta(t, Y.) \geqslant r_t \geqslant 0, \;\; \mathbb{P}^\pi - a.s. \tag{6.8.3}$$

The functionals $\nu, \theta, \lambda, \kappa, \rho$ are bounded and continuous in the sense of Assumption 6.3.1.

Assumption 6.8.2. *The reference point G is a constant.*

Remark 6.8.2. The assumption concerning the coefficients ν and κ can be weakened, indeed, as it was pointed out in the last section, it is possible to obtain the same growth estimates for $\|A_t\|$ in this case. This allows to include a myriad of stochastic volatility models and mean reverting economic factors. On the other hand, the assumption concerning the boundedness of ρ, θ and λ cannot be relaxed.

As in Subsection 6.4, we consider a relaxed setting. With this purpose in mind, we define $\theta^r(t, y.) = \theta(t, y.) - r_t$. In what follows, E is the 2×2 matrix such that $E^{11} = 1$ and $E^{ij} = 0$ otherwise.

Definition 6.8.3. We define the following family of sets.

$$A_t(x., y.) = \left\{ (a,b) \in \mathbb{A} \,\middle|\, a = \frac{1}{2}\kappa^2(t, y.)\, E + \frac{1}{2} m \lambda^2(t, y.)\, x_t^2 \begin{pmatrix} \rho^2(t, x.) & \rho(t, x.) \\ \rho(t, x.) & 1 \end{pmatrix}, \right. \tag{6.8.4}$$

$$\left. b = \begin{pmatrix} \nu(t, y.) \\ 0 \end{pmatrix} + (l x_t \theta^r(t, y.) + r_t x_t) \begin{pmatrix} \rho(t, x.) \\ 1 \end{pmatrix}, \quad \begin{matrix} 0 \leqslant m \leqslant 1 \\ 0 \leqslant l \leqslant \sqrt{m} \end{matrix} \right\}. \tag{6.8.5}$$

The following lemma enables us to use results in chapter 5, Theorem 5.4.3, originally proved in [32].

Lemma 6.8.4. *The set $A_t(x.,y.)$ is closed, convex and bounded for each $(x.,y.) \in \mathcal{W}_T^2$ and each $t \in [0,T]$.*

Proof. Only convexity needs to be checked. Let $0 \leqslant \mu \leqslant 1$ and (a,b), $(\alpha,\beta) \in A_t(x.,y.)$ then the convex linear combination $\mu a + (1-\mu)\alpha$ is equal to

$$\frac{1}{2}\kappa^2(t,y.)\,E_1 1 + \frac{1}{2}\left(\mu m + (1-\mu)\,m'\right)\lambda^2(t,y.)\,x_t^2\begin{pmatrix} \rho^2(t,x.) & \rho(t,x.) \\ \rho(t,x.) & 1 \end{pmatrix},$$

and $\mu b + (1-\mu)\beta$ equals

$$\begin{pmatrix} \nu(t,y.) \\ 0 \end{pmatrix} + \left(\left(\lambda l + (1-\lambda)\,l'\right)x_t\theta^r(t,y.) + r_t x_t\right)\begin{pmatrix} \rho(t,x.) \\ 1 \end{pmatrix}.$$

As $\mu l + (1-\mu)l' \leqslant \mu\sqrt{m} + (1-\mu)\sqrt{m'} \leqslant \sqrt{\mu m + (1-\mu)m'}$ so $\mu(a,b) + (1-\mu)(\alpha,\beta) \in A_t(x.,y.)$. □

The estimates of Lemma 6.5.1 apply to this case as well, by Assumption 6.8.1 and Assumption 6.8.2 one obtains a similar condition on A_t

$$\|A_t(\zeta.)\| \leqslant K\left(1 + |X_t|^2\right).$$

This allows to use the results of Theorem 5.4.3, just as they were used in the last section; using the class of relaxed controls defined below.

Definition 6.8.5. We say that $\overline{\pi} \in \overline{\Pi}$ if

$$\pi := \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, X_t, Y_t, (B_t, W_t), (x,y)\right)$$

with

(a) $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ a complete filtered probability space whose filtration satisfies the usual conditions;

(b) the two-dimensional process $\xi_t := (B_t, W_t)$ is a standard $\mathcal{F}_t$-Wiener process;

(c) the vector $(x,y) \in (0,\infty) \times \mathbb{R}$ is the initial endowment of the portfolio process X_t and the initial state of the economic factors Y_t, respectively;

(d) there exists an $\mathbb{A}$-valued, $\mathcal{F} \otimes \mathcal{B}([0,T])$ measurable and $\mathcal{F}_t$-adapted process denoted by (a_t, b_t) such that

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} + \int_0^t \sqrt{2a_s}\,d\xi_s + \int_0^t b_s ds \tag{6.8.6}$$

(e) For almost all $(\omega,t) \in \Omega \times [0,T]$, we have $(a_t,b_t) \in A_t(X.,Y.)$ (i.e. we can choose a pair (m_t, l_t) in a "measurable way").

As explained before, the vectorial form of the equations (6.2.4) and (6.2.5) can be rewritten. Define

$$\sigma_t := \begin{pmatrix} \kappa(t, y.) & \rho(t, x.)\sqrt{m_t} x_t \lambda(t, y.) \\ 0 & \sqrt{m_t}\lambda(t, y.)\, x_t \end{pmatrix}$$

and the drift is given by

$$b_t = \begin{pmatrix} \nu(t, y.) \\ 0 \end{pmatrix} + (x_t l_t \theta^r(t, Y.) + x_t r_t) \cdot \begin{pmatrix} \rho(t, x.) \\ 1 \end{pmatrix}$$

Given a relaxed control $\overline{\pi}$ the processes X_t, Y_t are $\mathcal{F} \otimes \mathcal{B}([0, T])$-measurable and $\mathcal{F}_t$-adapted such that for all $t \in [0, T]$

$$dY_t = \nu(t, Y.)\, dt + \kappa(t, Y.)\, dB_t + \rho(t, X.)\, dX_t, \quad Y_0 = y, \tag{6.8.7}$$

$$dX_t = [l_t(\theta(t, Y.) - r_t)\, X_t + r_t \cdot X_t]\, dt + \sqrt{m_t}\lambda(t, Y.)\, X_t dW_t, \quad X_0 = x. \tag{6.8.8}$$

The proof of the following result is based on that of Theorem 6.3.1.

Main Theorem 6.8.6. *Let Assumptions 6.3.2, 6.3.3, 6.8.1 and 6.8.2 be in force then the problem is well-posed and $\Pi' \neq \varnothing$ (the identically zero strategy belonging to Π', where Π' is defined analogously to Subsection 6.3). There is a $\hat{\pi} \in \Pi'$ such that the supremum in (6.2.9) is attained.* □

Remark 6.8.7. The assumption made on the factors Y_t to be one-dimensional is stated to simplify computations, in case $\nu(t, y.)$ is $\mathbb{R}^d-$valued and $\kappa(t, y.)$ is $\mathbb{R}^{d \times d_1}-$valued, Definitions 6.8.1, 6.8.4 and 6.8.5 are changed accordingly. Moreover, the estimates in Proposition 6.5.3, or in (6.5.6) and (6.5.7) hold true with no significant changes, hence relatively compactness of the laws of $(X., Y.)$ holds.

8.1 Auxiliary lemmas

Some proofs of auxiliary results are included in this section.

Proof of Proposition 6.5.2. We shall write $\xi_s = (W_s, B_s)$. Suppose $m \geqslant 2$. The notation $|\cdot|$ will be used to denote Euclidean norm in spaces of various dimensions. By a localisation argument, we can assume that $\sup_{t \in [0,T]} |X_t|^m \leqslant N$, for a positive constant $N > 0$. Then

$$|\zeta_t|^m = \left[(X_t)^2 + |Y_t|^2\right]^{\frac{m}{2}} \leqslant 2^{\frac{m}{2}-1} \cdot [|X_t|^m + |Y_t|^m], \tag{6.8.9}$$

so it is enough to obtain that the moments of each of the processes Y_t and X_t satisfy (6.5.4). Set $b_{t,2} = (l_t \theta_t(Y.)\, X_t)$ and $\sigma_{t,2} = (0, m_t \lambda_t(Y.)\, X_t)$

$$\mathbf{E}^\pi\left[\sup_{t \leqslant T} |X_t|^m\right] \leqslant 3^{m-1}\left(|X_0|^m + \mathbf{E}^\pi\left(\int_0^T |b_{s,2}| ds\right)^m + \mathbf{E}^\pi\left[\sup_{t \leqslant T}\left|\int_0^t \sigma_{s,2} d\xi_s\right|^m\right]\right). \tag{6.8.10}$$

By Jensen's inequality and by Burkholder-Davis-Gundy inequality

$$\mathbf{E}^\pi\left[\sup_{t \leqslant T} |X_t|^m\right] \leqslant 3^{m-1}\left(|X_0|^m + \mathbf{E}^\pi\left(\int_0^T |b_{s,2}| ds\right)^m + C_m \mathbf{E}^\pi\left[\left|\int_0^T |\sigma_{s,2}|^2 ds\right|^{\frac{m}{2}}\right]\right),$$

for some $C_m > 0$.

We can apply again Jensen's inequality (now with respect the "uniform density" on $[0, T]$)

$$\left(T\int_0^T \frac{|b_{s,2}|}{T} ds\right)^m \leqslant T^{m-1} \cdot \int_0^T |b_{s,2}|^m ds, \text{ and } \left|T\int_0^T \frac{\|\sigma_{s,2}\|^2}{T} ds\right|^{\frac{m}{2}} \leqslant T^{m/2-1} \cdot \int_0^T \|\sigma_{s,2}\|^m ds, \tag{6.8.11}$$

here $\|\sigma_{s,2}\|^m = m_t^m |\lambda_t(Y.)|^m \cdot X_t^m$ and $|b_{s,2}|^m = l_t^m \theta_t^m(Y.) X_t^m$, however, we can use the estimate (6.5.1) above,

$$\mathbf{E}^\pi\left[\sup_{t\leqslant T} |X_t|^m\right] \leqslant 3^{m-1}\left(\mathbf{E}^\pi |X_0|^m + K(m,T)\cdot \mathbf{E}^\pi\left[\int_0^T \left(1+\sup_{t\leqslant s}\|\zeta_t\|\right)^m ds\right]\right),$$

similarly,

$$\mathbf{E}^\pi\left[\sup_{t\leqslant T} |Y_t|^m\right] \leqslant 3^{m-1}\left(\mathbf{E}^\pi |Y_0|^m + K'(m,T)\mathbf{E}^\pi\left[\int_0^T \left(1+\sup_{t\leqslant s}\|\zeta_t\|\right)^m ds\right]\right),$$

for constants K, K' depending on m and the fixed finite horizon $T > 0$. Then we have

$$\mathbf{E}^\pi\left[\sup_{t\leqslant T} \|\zeta_t\|^m\right] \leqslant K(m)\left(\mathbf{E}^\pi \|\zeta_0\|^m + \left[\int_0^T 1+\mathbf{E}^\pi\left(\sup_{t\leqslant s}\|\zeta_t\|\right)^m ds\right]\right), \tag{6.8.12}$$

for some $K(m,T) > 0$ so by Grömwall's inequality,

$$\mathbf{E}^\pi\left[\sup_{t\in[0,T]} \|\zeta_t\|^m\right] \leqslant L(m,T),$$

with a fixed constant $L(m)$, not depending on N, for all $\pi \in \Pi$. The general case follows by monotone convergence.
The case $0 < m < 2$ follows from the monotonicity of the norms. □

Proof of Proposition 6.5.3. For simplicity, the processes $Y_t^{\overline{\pi}}$ and $X_t^{\overline{\pi}}$ are denoted by Y_t, X_t respectively, the notations B, W are thus also self-explanatory. We follow the same arguments that were used in Chapter 5, in the proof of Theorem 5.4.3.

As in Proposition 6.5.2, it is enough to show a similar estimate to (6.5.5) for each of the coordinates X_t and Y_t

$$\mathbf{E}^\pi |Y_t - Y_s|^\eta \leqslant K_1 |t-s|^{\eta/2} \text{ and } \mathbf{E}^\pi |X_t - X_s|^\eta \leqslant K_2 |t-s|^{\eta/2}. \tag{6.8.13}$$

By Assumption 6.3.1, and by moments estimates of stochastic integrals, see Corollary 2.5.3 in [36] (or Burkhölder-Davis-Gundy) inequality and Jensen's inequality

$$\mathbf{E}^\pi |Y_t - Y_s|^\eta \leqslant 2^{\eta-1}\cdot\left[\mathbf{E}^\pi\left(\int_s^t |\nu(Y.)|\, dr\right)^\eta + \mathbf{E}^\pi\left(\left|\int_s^r \kappa_r(Y.)\, dB_r\right|^\eta\right)\right],$$

$$\mathbf{E}^\pi |Y_t - Y_s|^\eta \leqslant 2^{\eta-1}\cdot\left[M^\eta |t-s|^\eta + C_\eta \cdot \mathbf{E}^\pi\left(\int_s^t \kappa_r^2(Y.)\, dr\right)^{\eta/2}\right],$$

thus

$$\mathbf{E}^{\pi}\left|Y_t-Y_s\right|^{\eta}\leqslant 2^{\eta-1}\cdot\left[M^{\eta}\left|t-s\right|^{\eta}+C_{\eta}M^{\eta}\left|t-s\right|^{\eta/2}\right]\leqslant 2^{\eta-1}\cdot\left[M^{\eta}T^{\eta/2}+C_{\eta}M^{\eta}\right]\cdot\left|t-s\right|^{\eta/2}.$$

The second estimate relies upon Proposition 6.5.2 or in moments estimates of stochastic integral, see [36], Corollary 2.5.3:

$$\mathbf{E}^{\pi}\left|X_t-X_s\right|^{\eta}\leqslant 2^{\eta-1}\cdot\left[\mathbf{E}^{\pi}\left(\int_s^t\left|l_r\theta_r\left(Y_\cdot\right)X_r\right|dr\right)^{\eta}+\mathbf{E}^{\pi}\left(\left|\int_s^t\lambda_r\left(Y_\cdot\right)\sqrt{m_r}\,X_r dW_r\right|^{\eta}\right)\right],$$

$$\leqslant 2^{\eta-1}\cdot M^{\eta}\cdot\left[\mathbf{E}^{\pi}\left(\int_s^t\left|X_r\right|dr\right)^{\eta}+C_{\eta}\mathbf{E}^{\pi}\left(\left|\int_s^t\left|X_r\right|^2dr\right|^{\eta/2}\right)\right],$$

$$\mathbf{E}^{\pi}\left|X_t-X_s\right|^{\eta}\leqslant 2^{\eta-1}\cdot M^{\eta}\left[(t-s)^{\eta}\cdot\mathcal{N}\left(\eta,T\right)+(t-s)^{\eta/2}\cdot\mathcal{N}\left(\eta,T\right)\right]=K_3\left|t-s\right|^{\eta/2},$$

where $K_3=2^{\eta-1}M^{\eta}\mathcal{N}\left(\eta,T\right)\cdot\left(T^{\eta/2}+1\right)$ and $\mathcal{N}\left(\eta,T\right)$ is the upper bound of $\sup_{\overline{\pi}\in\overline{\Pi}}\mathbf{E}^{\pi}\left[\sup_{t\leqslant T}\left|X_t\right|^{\eta}\right]$ in Proposition 6.5.2. Note that the constants do not depend on $\overline{\pi}$ as neither M nor $\mathcal{N}\left(\eta,T\right)$ do. $\square$

In this chapter we have provided a generalization of results in [41] and [45]. We do not use any kind of completeness whatsoever, indeed, our results in Theorem 6.3.1 and Theorem 6.8.6 depend on facts concerning compactness of laws of continuous semimartingales and the martingale approach to Itô Stochastic Differential Equations. As noted before, an advantage of this approach is provided in terms on the conditions we impose to the market, no Lipschitz continuity on the coefficients is assumed. However, as it happens with the martingale problem formulation, in order to ensure well-posedness of the problem, boundedness is imposed for the sake of simplicity, although this is not essential; our condition can be replaced by a linear growth condition on the economic factor $\{Y_t\}_{t\in[0,T]}$, and the estimates hold true in this case as well. On the other hand questions concerning analytical approximations of optimal strategies remain open.

The relaxation technique is common in many other problems in Mathematics and in Control theory. It is interesting to note how this fits with the compactness principle proved in [32]. Finally we would like to remark that our concept of relaxation, is different to the so-called 'relaxed' control approach in stochastic control theory, see [23] or [8].

Chapter 7

Conclusions and future research

As explained throughout this work, we analysed the problem of optimal investment when the principles of CPT are described by 'behavioural functionals' defined in chapter 3 and chapter 6. We provided verifiable conditions that ensure well-posedness and showed existence of optimal strategies, we used properties of martingale measures (in the discrete-time setting) or topological properties of the set of laws (continuous-time).

It is shown how topics such as utility maximisation; martingale problems; weak convergence of processes; are relevant to our main results, from a mathematical point of view. The reason for using these methods are justified by the nature of the problem; the limitations imposed by other methods such as dynamic programming or duality methods and the interest in general and verifiable conditions. The drawback of this approach is directly related to the limitations of several existence proofs that are non constructive. We consider that our approach could be described as "probabilistic".

We wish to stress that despite of the possible limitations of our approach, our main theorems generalise many of the existing results in the literature such as those in [11], [41] and [46]. An important question that is left unanswered in the main theorems in Chapter 6 is the fact that the relaxation seems to work in the case of a single risky asset. Whether this can be easily extended to the multi-asset case, using the same idea, is a compelling research problem.

Another important question that remains unsolved is the uniqueness of the optimal law. It seems that our methods do not enable us to conclude uniqueness, and some examples given in [11], [41] and [37] seem to suggest that, in general, uniqueness may not hold.

An interesting feature of our approach in Chapter 6 is the following: our results are based upon stochastic control techniques and do not rely on any assumptions related to no arbitrage. The idea of relaxation may have the potential of being adapted to many other problems and clearly, the limitations of our methods are linked to the assumptions 5.4.1 and 5.2.1 and to the requirements of boundedness and convexity of the sets A_t. This is another natural direction for future research.

Finally, questions concerning approximation procedures to obtain so called $\varepsilon-$optimal strategies, remain another research direction.

Appendix A

Regular conditional probability

1 Definitions

In this appendix we review some of the basic definitions concerning the existence of a regular conditional probability. Let $(E_i, \mathcal{E}_i)$ for $i = 1, 2$ two measurable spaces.

Definition A.1.1. We say that $N : E_1 \times \mathcal{E}_2 \to \overline{\mathbb{R}}_+$ is a *kernel* from $(E_1, \mathcal{E}_1)$ to $(E_2, \mathcal{E}_2)$ to a function such that

a) For all $x \in E_1$, $N(x, \cdot)$ is a finite measure on $\mathcal{E}_2$,

b) For all $A \in \mathcal{E}_2$, $N(\cdot, A)$ is a $\mathcal{E}_1-$measurable.

We say that the kernel is a transition probability if for all $x \in E_1$ the measure $N(x, \cdot)$ is a probability measure.

2 Main results

Theorem A.2.1. *Let $(\Omega, \mathcal{F})$ be a measurable space, $(E, \mathcal{E})$ be a Polish space and $\mathbb{P}$ be a probability measure on $\mathcal{F} \otimes \mathcal{E}$. Denote by $\pi_1 : \Omega \times E \to E$ the projection over Ω, $\pi_1((\omega, x)) = \omega$ and $P_1 = \mathbb{P} \circ \pi_1^{-1}$ then there exists a transition probability $Q_1 : \Omega \times \mathcal{E} \to [0, 1]$ from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ such that $\mathbb{P}$ is disintegrated with respect to P_1.*

We include for the sake of completeness the following important theorem. A proof can be found in [21].

Theorem A.2.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}$ be sub σ-algebras. Let $(E, \mathcal{E})$ be a Borel space and $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ be a random variable taking values in E. Then there exists a regular conditional probability given $\mathcal{G}$. This means that there is a mapping $Q(\cdot, \cdot) : \Omega \times \mathcal{G}_1 \to [0, 1]$ with the following properties*

(a) *For each $\omega \in \Omega$, $Q(\omega, \cdot)$ is a probability measure on $\mathcal{G}_1$*

(b) *For each $\mathcal{G}_1$ measurable set A_1, the r.v. given by $Q(\omega, A_1)$ is $\mathcal{F}_2-$measurable.*[1]

(c) *$\mathbb{P} -$ a.s. the mapping $Q(\omega, A_1)$ is a version of the conditional probability of A_1 given $\mathcal{F}_2$, i.e. $Q(\omega, A_1) = \mathbb{P}\left(A_1 \middle| \mathcal{F}_2\right)$.*

[1]When such a mapping satisfies conditions (a) and (b) we say that Q is a transition probability given $\mathcal{F}_2$

Appendix B

Essential infimum and essential supremum

1 Main result

A review of the definition of the essential infimum and the essential supremum. For further properties and proofs see [38].

Proposition B.1.1. *Let $\{X_i\}_{i\in I}$ be any sequence of real-valued random variables defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. There exist random variables, denoted by* $\operatorname{ess\,inf}_{i\in I} X_i$ *and* $\operatorname{ess\,sup}_{i\in I} X_i$*, such that*

- *For any $i \in I$ X_i,* $\operatorname{ess\,inf}_{i\in I} X_i \leqslant X_i$ $\mathbb{P}-a.s.$ *and* $X_i \leqslant \operatorname{ess\,sup}_{i\in I} X_i$ $\mathbb{P}-a.s.$
- *If W, Z are other random variables such that for all $i \in I$ $X_i \leqslant Z$ $\mathbb{P}-a.s.$ or $W \leqslant X_i$ $\mathbb{P}-a.s.$ then*

$$\operatorname*{ess\,sup}_{i\in I} X_i \leqslant Z \quad and \quad W \leqslant \operatorname*{ess\,inf}_{i\in I} X_i \ \mathbb{P}-a.s.$$

2 A compactness principle in $L^0\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)$

The next lemma is used in Section 2, it is stated as a (relatively) compactness principle in $L^0\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)$. It is proved in [29] and we include it here for ease of reference.

Lemma B.2.1. *Let $\eta_k \in L^0\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)$ be such that $\overline{\eta} := \liminf |\eta_k| < \infty$. Then there are $\tilde{\eta}^k \in L^0\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)$ for all ω the sequence $\tilde{\eta}^k(\omega)$ is a convergent subsequence of $\eta^n(\omega)$.*

Appendix C

Set-valued functions

1 Definitions

This appendix contains relevant definitions and results concerning the measurability of set-valued functions. Intuitively, a set-valued function establishes a correspondence between a point in a measurable space $(X, \mathcal{F})$ and a measurable set in another measurable space (here this measurable space shall be $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$. We review the concept of a measurable set-valued function, for a thorough revision of this topic see [48] or [2].

Definition C.1.1. Assume $(\Omega, \mathcal{F})$ is a measurable space. Let $S : \Omega \rightrightarrows \mathbb{R}^d$ be a set-valued function (we use this notation to indicate that S is not a function but a multi-valued function). We say that that the multi-valued function (or set-valued function) S is measurable if for any $D \in \mathcal{B}\left(\mathbb{R}^d\right)$

$$S^{-1}(D) := \{\omega \in \Omega \,|\, S(\omega) \cap D \neq \varnothing\} \in \mathcal{F},$$

The graph of a set-valued function S is defined to be the set

$$gph(S) = \left\{(\omega, x) \in \Omega \times \mathbb{R}^d : \ \omega \in \operatorname{dom} S, \ x \in S(\omega)\right\}.$$

The domain of S is the set $S^{-1}\left(\mathbb{R}^d\right) = \{\omega \in \Omega \, : \, S(\omega) \neq \varnothing\}$.

Definition C.1.2. We say that a set-valued function is closed if $S(\omega)$ is closed for any ω. A set-valued function $\mathscr{H} : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is osc (*outer semi-continuous*) at a point $\overline{x}$ if

$$\lim_{x \to \overline{x}} \mathscr{H}(x) \subset \mathscr{H}(\overline{x}),$$

where the set $\limsup_{x \to \overline{x}} \mathscr{H}(x)$ is defined below

$$\limsup_{x \to \overline{x}} \mathscr{H}(x) = \{u : \exists x^\nu \to \overline{x} \ \text{ and } \ u^\nu \to u \ \text{ with } \ u^\nu \in \mathscr{H}(x^\nu)\}.$$

2 Auxiliary results

The following lemma simplifies the verification of measurability of a set-valued function, but first we state Lemma C.2.1, in turn this simplify the verification of the osc property.

Lemma C.2.1. *A set-valued function $S : \mathbb{R}^d \rightrightarrows \mathbb{R}^m$ is osc if and only if gphS is closed in $\mathbb{R}^d \times \mathbb{R}^m$.*

We have an important characterisation of the measurability of set-valued functions. This is Theorem 14.8 p. 648 in [48].

Theorem C.2.2. *Suppose $S : X \rightrightarrows \mathbb{R}^d$ is a set-valued function. The following statements are equivalent*

a) *The set-valued function S is measurable.*

b) *The graph of the function $gph\,(S)$ is a measurable set of $\mathcal{F} \otimes \mathcal{B}\left(\mathbb{R}^d\right)$.*

c) *$S^{-1}(D) \in \mathcal{F}$ for all sets $D \in \mathcal{B}\left(\mathbb{R}^d\right)$*

We use these results in the context of chapter 2. Suppose $\mathcal{F}_0$ is complete with respect to $\mathbb{P}$. By definition of the support of the measure $\mu(\omega, \cdot) = \mathbb{P}\left(\Delta S_t \in \cdot \,\middle|\mathcal{F}_{t-1}\right)(\omega)$, the set $supp\,(\mu(\omega, \cdot))$ is closed and there is a version such that $\mu(\omega, \cdot)$ is a probability measure, denote by $\overline{\Omega}$ the set having full measure on which this hold. Define the set-valued function $\mathscr{S} : \Omega \rightrightarrows \mathbb{R}^d$

$$\mathscr{S}(\omega) := \begin{cases} supp\,(\mu(\omega, \cdot)) & \omega \in \overline{\Omega} \\ \mathbb{R}^d & \omega \in \Omega/\overline{\Omega}, \end{cases} \tag{C.2.1}$$

Proposition C.2.3. *The multi-valued function $\mathscr{S} : \Omega \rightrightarrows \mathbb{R}^d$ is $\mathcal{F}_{t-1}-$measurable.*

Proof. By properties of the regular conditional probability measure (see Theorem A.2.2) we know that for any $B \in \mathcal{B}(\mathbb{R}^d)$, $\mathbb{P}\left(\Delta S_t \in B\middle|\mathcal{F}_{t-1}\right)(\omega)$ is a $\mathcal{F}_{t-1}-$measurable r.v. for all $\omega \in \overline{\Omega} \subset \Omega$ and it has full measure. Thus, for $\overline{\Omega} \cap \left[\mathbb{P}\left(\Delta S_t \in B\middle|\mathcal{F}_{t-1}\right)\right]^{-1}((0,1]) \in \mathcal{F}_{t-1}$ and $\overline{\Omega} \cap \left[\mathbb{P}\left(\Delta S_t \in B\middle|\mathcal{F}_{t-1}\right)\right]^{-1}((0,1]) \times B \in \mathcal{F}_{t-1} \otimes \mathcal{B}\left(\mathbb{R}^d\right)$. To simplify notation we write $\mu(\omega, \cdot) = \mathbb{P}\left(\Delta S_t \in \cdot\,\middle|\mathcal{F}_{t-1}\right)$.

We use the theorem C.2.2 (Theorem 14.8 in [48]) on the equivalence of the measurability of $\mathscr{S}(\omega)$. Hence we show that the graph of the set-valued function is $\mathcal{F}_{t-1} \otimes \mathcal{B}\left(\mathbb{R}^d\right)-$measurable. This follows from the equality given below

$$gph\,(\mathscr{S}) = \bigcap_{\rho \in \mathbb{Q}^+} \bigcup_{q \in \mathbb{Q}^d} \left[\overline{\Omega} \cap \left(\mu\left(B\left(q, \rho\right)\right)^{-1}((0,1])\right) \times B\left(q, \rho\right)\right] \bigcup \left(\Omega \backslash \overline{\Omega}\right) \times \mathbb{R}^d, \tag{C.2.2}$$

the identity holds, as the vector space $\mathbb{R}^d$ is separable. We explain below why C.2.2 holds
Let $\omega \in \overline{\Omega}$ and $x \in \mathscr{S}(\omega)$, by definition of the support for all $d \in \mathbb{Q}^+$ we have $\mu(\omega, B(x, d)) > 0$, as $\left\{B(q, \rho) : q \in \mathbb{Q}^d, \rho \in \mathbb{Q}^+\right\}$ is a countable basis, there is a ball $B(q, \rho)$ such that $x \in B(q, \rho)$ for some $(q, \rho) \in \mathbb{Q}^d \times \mathbb{Q}^+$, and then $\mu(\omega, B(q, \rho)) > 0$, thus $(\omega, x) \in \bigcap_{\rho \in \mathbb{Q}^+} \bigcup_q \left[\overline{\Omega} \cap \left(\mu\left(B\left(q, \rho\right)\right)^{-1} ((0,1])\right) \times B(q, \rho)\right]$, in the case $\omega \in \Omega/\overline{\Omega}$ the inclusion follows trivially.

If $(\omega, x) \in \bigcap_{\rho \in \mathbb{Q}^+} \bigcup_q \left[\overline{\Omega} \cap \left(\mu\left(B\left(q, \rho\right)\right)^{-1}((0,1])\right) \times B(q, \rho)\right] \bigcup \left(\Omega \backslash \overline{\Omega}\right) \times \mathbb{R}^d$, let V be an open neighbourhood of $x \in V$, there are $\{B(q_i, r_i)\}_{i \geqslant 1}$ such that $V = \bigcup_i B(q_i, r_i)$ thus, there is i such that $x \in B(q_i, r_i)$ there is, $p \in \mathbb{Q}^d$ such that $x \in B\left(p, \frac{r_i}{2}\right)$ and $\mu\left(\omega, B\left(p, \frac{r_i}{2}\right)\right) > 0$ taking $\delta \in \mathbb{Q}^+$ small enough (by the base property) we can find $x \in B(q, \delta) \subset B\left(p, \frac{r_i}{2}\right) \cap B(q_i, r_i)$ but then $\mu(\omega, V) \geqslant \mu(\omega, B(q_i, r_i)) > 0$ therefore $x \in supp\,(\mu(\omega, \cdot))$. In other words, $(\omega, x) \in \mathscr{S}(\omega)$.

As the measure $\mu(\omega, \cdot)$ is a regular conditional probability given $\mathcal{F}_{t-1}$, for $\omega \in \overline{\Omega}$, $\mu(\omega, B(q, \rho))^{-1}((0,1]) \in \mathcal{F}_{t-1}$ then the set in (C.2.2) is $\mathcal{F}_{t-1} \otimes \mathcal{B}\left(\mathbb{R}^d\right)-$measurable and then $\mathscr{S}$ is measurable. □

Proposition C.2.4. *If $\mathscr{G} : \Omega \rightrightarrows \mathbb{R}^d$ is a measurable set-valued function and then $aff(\mathscr{G}(\omega))$ (the affine subspace of the set $G(\omega)$) is also measurable.*

This is exercise 14.12 d) in [48].

Remark C.2.5. If $A \in \mathcal{F}_{t-1} \otimes \mathcal{B}(\mathbb{R}^d)$ and $Y : \Omega \to \mathbb{R}^d$ is Borel measurable then the set $\{\omega \in \Omega : (\omega, Y(\omega)) \in A\} \in \mathcal{F}_{t-1}$.

Indeed, this remark holds, as we have a composition of two measurable functions $\mathbb{I}_A(\omega, x) \circ (id_\Omega, Y)$.

Proposition C.2.6. *Let $\mu(\omega, \cdot) = \mathbb{P}(\Delta S_t \in \cdot | \mathcal{F}_{t-1})(\omega)$ a regular conditional probability given $\mathcal{F}_{t-1}$ and $\mathscr{D}_t(\omega) := aff(\mathscr{S}(\omega))$ then $\mathbb{P}((\omega, \Delta S_t(\omega)) \in \mathscr{D}_t(\omega)) = 1$.*

This is an application of Theorem III-44 in [21] or a consequence of the following identity

$$\mathbb{P}\Big(\omega \in \Omega : (\omega, \Delta S_t(\omega)) \in G_t\Big) = \int_{\overline{\Omega}} \mathbb{P}\Big(\pi_\omega G_t \big| \mathcal{F}_{t-1}\Big)(\omega)\, d\mathbb{P}(\omega) = 1, \tag{C.2.3}$$

where $G_t = \left\{(\omega, y) \in \Omega \times \mathbb{R}^d : y \in \mathscr{D}_t(\omega)\right\} \in \mathcal{F}_{t-1} \otimes \mathcal{B}(\mathbb{R}^d)$ by Proposition C.2.4 and Theorem C.2.2. The equality in C.2.3 is a consequence of the following facts, $\overline{\Omega}$ has full measure, Fubini's theorem applied to the regular conditional probability $\mu(\omega, \cdot)$ and the definition of the support of a measure.
We recall the following lemma on composition of set-valued functions, we are following [48].

Proposition C.2.7. *Let $(\Omega, \mathcal{G})$ be a measurable space and let $S : \Omega \rightrightarrows \mathbb{R}^n$ be a closed-valued and $\mathcal{G}-$measurable set-valued function. For each $\omega \in \Omega$ suppose $M(\omega, \cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is an osc set-valued function. Suppose that the 'graphical mapping' $\omega \longmapsto gph(M(\omega, \cdot))$ is $\mathcal{G}-$measurable, then $\omega \longmapsto M(\omega, S(\omega))$ is $\mathcal{G}-$measurable.*

Let $(x_i)_{i \leqslant n} \subset \mathbb{R}^d$ and denote by $\mathscr{L}(x_1, x_2, \ldots, x_n)$ the linear span of the vectors $\{x_i\}_{i \leqslant n}$. An application of measurable selection theorems yield the following result. For a proof, refer to [21] or [48].

Lemma C.2.8. *Suppose that $\mathscr{D}(\omega) \neq \{0\}$ for any $\omega \in \Omega$. Then there is a measurable set-valued function $\sigma : \Omega \to \Pi_{i=1}^d \mathbb{R}^d$ such that $\sigma(\omega) = (\sigma_i(\omega))_{i \leqslant d}$ and $\mathscr{L}(\sigma_1, \sigma_2, \ldots, \sigma_d) = \mathscr{D}(\omega)$.*

From this lemma we have that

$$\{(\omega, x) : x \in \mathscr{D}(\omega), \langle x - \xi(\omega), \sigma_i(\omega)\rangle = 0\}, \tag{C.2.4}$$

is measurable and $\{(\omega, x) : x \in \mathscr{D}(\omega), \langle x - \xi(\omega), \sigma_i(\omega)\rangle = 0\} = \Big\{(\omega, x) : \omega \in \Omega,\, x = \hat{\xi}(\omega)\Big\} = gph\Big(P_{\mathscr{D}_t(\omega)} \circ \xi\Big)$. In other words, the graph of the set-valued function $P_{\mathscr{D}_t(\omega)} \circ \xi$ is measurable (notice that this multi-valued function maps ω to the singleton set $\Big\{\hat{\xi}(\omega)\Big\}$). This implies that the r.v. $\hat{\xi}(\omega)$ is $\mathcal{F}_{t-1}$.

An alternative to the last proof (without using measurable selection theorems).

Lemma C.2.9. *Suppose $D_t(\omega)$ is defined as above (the image of the set-valued defined by $\mathscr{D}_t(\omega)$). Let $\xi : \Omega \to \mathbb{R}^d$ be a random variable and denote by $\hat{\xi}(\omega)$ its orthogonal projection into $D_t(\omega)$. If $\xi \in \Xi_{t-1}$ then $\hat{\xi} \in \Xi_{t-1}$.*

Proof. We start by noticing that Proposition C.2.6 shows that $\mathscr{D}_t$ is a measurable closed and set-valued function. By Theorem 14.3 (j) in [48] the function $d_t : \Omega \to \mathbb{R}_+$ defined by $d_t(\omega) := d(x, D_t(\omega))$ is measurable for each x.
Denoting by $P_{D_t(\omega)}$ the orthogonal projection to $D_t(\omega)$ (for each $\omega \in \Omega$), $P_{D_t(\omega)} \in L(\mathbb{R}^d; \mathbb{R}^d)$.
Let us define by $\mathscr{P}(\omega)$ the set-valued function that associates to each ω the graph of $P_{D_t(\omega)}$, i.e. $\mathscr{P}(\omega) = gph\left(P_{D_t(\omega)}\right)$.
We claim that $gph(\mathscr{P}) \in \mathcal{F}_{t-1} \otimes \mathcal{B}\left(\mathbb{R}^d \times \mathbb{R}^d\right)$. Indeed,

$$gph\Big(\mathscr{P}\Big) = \left\{(\omega, x, y) \in \Omega \times \mathbb{R}^d \times \mathbb{R}^d \,:\, (x, y) \in \mathscr{P}(\omega)\right\} =$$

$$= \left\{(\omega, x, y) \,:\, y = P_{D_t(\omega)} x\right\} = \left\{(\omega, x, y) \,:\, |x - y| - d(x, D_t(\omega)) = 0\right\}, \tag{C.2.5}$$

as the function in (C.2.5), $|x - y| - d(x, D_t(\omega))$, is measurable. The claim on the graph of $\mathscr{P}$ follows. If $\omega \in \Omega$ is fixed them the 'set-valued' function $P_{D_t(\omega)} : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ given by the 'singleton set' $\left\{P_{D_t(\omega)} x\right\}$ and

$$gphP_{D_t(\omega)} x = \{(x, y) \,:\, |x - y| - d(x, D_t(\omega)) = 0\},$$

is closed. The set-valued function $S_\xi := \{\xi(\omega)\}$ is a closed valued (for each ω the image is a closed set) and measurable, as in this case for any U open

$$S^{-1}(U) = \left\{\omega : S^{-1}(\omega) \bigcap U \neq \varnothing\right\} = \left\{\omega : S^{-1}(\omega) \in U\right\} = \xi^{-1}(U) \in \mathcal{F}_{t-1}.$$

Then by proposition C.2.7 the set-valued function $P_{D_t(\omega)} \xi(\omega)$ is $\mathcal{F}_{t-1}$−measurable i.e. $\hat{\xi} \in \Xi_{t-1}$. □

Finally, we briefly explain proposition 4.6 in [44] in our context.

3 Lemma 2.4.18

Proposition C.3.1. *Let $\xi \in \Xi_{t-1}$, and as above, $\hat{\xi}(\omega)$, the orthogonal projection of ξ into $D_t(\omega)$. Furthermore,*

$$\mathbf{E}\left[V(x + \langle \xi, \Delta S_t \rangle) \,|\mathcal{F}_{t-1}\right] = \mathbf{E}\left[V\left(x + \left\langle \hat{\xi}, \Delta S_t \right\rangle\right) |\mathcal{F}_{t-1}\right], \tag{C.3.1}$$

$\mathbb{P} - a.s.$ for every $x \in \mathbb{R}$.

Proof. As lemma C.2.9 shows, $\hat{\xi} \in \Xi_{t-1}$, on the other hand $\{\Delta S_t \in D_t\} \subset \left\{\langle \xi, \Delta S_t \rangle = \left\langle \tilde{\xi}, \Delta S_t \right\rangle\right\}$ then

$$\mathbb{P}\left(\langle \xi, \Delta S_t \rangle = \left\langle \hat{\xi}, \Delta S_t \right\rangle |\mathcal{F}_{t-1}\right) \geqslant \mathbb{P}\left(\Delta S_t \in D_t \middle| \mathcal{F}_{t-1}\right) = 1, \tag{C.3.2}$$

the last equality is by definition of D_t and the definition of the support of a measure, see Proposition C.2.6. Taking expectations in (C.3.2) we have that $\left\langle \hat{\xi}, \Delta S_t \right\rangle = \langle \xi, \Delta S_t \rangle$ *a.s.*
This implies $V(x + \langle \xi, \Delta S_t \rangle) = V\left(x + \left\langle \hat{\xi}, \Delta S_t \right\rangle\right)$ $\mathbb{P}$−a.s. and (C.3.1) follows. □

Bibliography

[1] Aldous, D. Stopping times and tightness. *Ann. Prob.* 2, 335–340.

[2] Aliprantis, C., and Border, K. *Infinite Dimensional Analysis.* Springer-Verlag, 1994.

[3] Barlow, M., Émery, M., Knight, F., Song, S., and Yor, M. Autour d'un théoreme de Tsirelson sur des filtrations browniennes et non browniennes. *Séminaire de Probabilités*, XXXII, Lecture Notes in Math. 1686:264–305, Springer, Berlin, 1998.

[4] Bernoulli, B. *Theoriae Novae de Mensura Sortis.* Commentarii Academiae Scientiarum Imperialis Petropolitanae. Volume V., 1738. Translated by L. Sommer as "Exposition of a New Theory on the Measurement of Risk", *Econometrica*, 22:23–36, 1954.

[5] Berkelaar, A. B., Kouwenberg, R., and T. Post. Optimal portfolio choice under loss aversion. *Rev. Econ. Stat.*, 86:973–987, 2004.

[6] Billingsley, P. *Convergence of Probability Measures.* 2nd Edition, *Wiley*, 1999.

[7] Black, F. Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177–181, 1976.

[8] Borkar, V. S., *Optimal Control of Diffusion Processes.* Pitman Research Notes in Math. 203, Longman Scientific and Technical, Harlow, UK, 1989.

[9] Brossard, J. Deux notions équivalentes d'unicité en loi pour les équations différentielles stochastiques. *Séminaire de Probabilités de Strasbourg* XXXVII, Lecture Notes in Math., vol. 1832:246–250, 2003.

[10] Campi L., and Del Vigna M. Weak insider trading and behavioural finance. *SIAM J. Financial Mathematics*, 3:242–279, 2012.

[11] Carassus L., and Rásonyi M. On optimal investment for a behavioral investor in multiperiod incomplete markets. *To appear in Math. Finance,* 2013. Available at `http://arxiv.org/abs/1107.1617`

[12] Carlier G., and Dana R.-A.. Optimal demand for contingent claims when agents have law invariant utilities. *Math. Finance*, 21:169–201, 2011.

[13] Cherny, A. S. On the uniqueness in law and the pathwise uniqueness for stochastic differential equations. *Theory Probab. Appl.* 46, 3:406-418 (2002).

[14] Cherny, A., and Madan D.P. New Measures for Performance Evaluation. *Review of Financial Studies*, 22:2571–2606, 2009.

[15] Chow, Y.S.C., and Teicher, H. *Probability Theory: Independence and Exchangeability. 2nd Edition*, 1978. *Springer-Verlag*, .

[16] Dalang R., Morton A., and Willinger W. Equivalent martingale measures and arbitrage in stochastic securities market models. *Stochastics and Stochastic Reports* 29:185–201, (1990).

[17] Delbaen, F. A remark on Slutsky's theorem. *Séminaire de probabilités (Strasbourg)* 32:313-315, (1998)

[18] Delbaen, F., and Schachermayer, W. A general version of the fundamental theorem of asset pricing. *Math. Ann.* 300:463–520, (1994).

[19] Delbaen, F., and Schachermayer, W. The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Ann.* 312:215–250, (1998).

[20] Delbaen, F., Kabanov, Y., and Valkeila, S. Hedging under transaction costs in currency markets: a discrete-time model. *Mathematical Finance*, 12:45-61, 2002.

[21] Dellacherie, C., and Meyer, P. A. *Probability and potentials.* 2nd Edition, *North-Holland*, 1980.

[22] Dudley, R. *Probability and Real analysis.* 2nd Edition, *Cambridge*, 1999.

[23] El Karoui, N., Nguyen, D., Jeanblanc-Picqué, M. *Compactification methods in the control of degenerate diffusions: existence of an optimal control*, *Stochastics and stochastic reports* 20, 3:169-219, (1987). Stochastics

[24] Gilboa, I., *Rational Choice.* MIT press (2010).

[25] Ikeda, N., and Watanabe, S. *Stochastic Differential Equations and Diffusion Processes. North-Holland* (1989).

[26] Jacod, J., and Shiryaev, A.N. *Limit theorems for stochastic processes. Springer-Verlag*, (1987).

[27] Jacod, J., and Shiryaev, A. N. Local martingales and the fundamental asset pricing theorems in the discrete-time case. *Finance Stoch.* 2:259–273,(1998).

[28] Jin, H. and Zhou, X. Y. Behavioural portfolio selection in continuous time. *Math. Finance*, 18:385–426, 2008.

[29] Kabanov, Yu., and Stricker, C. A teachers' note's on no arbitrage criteria. *Séminaire de Probabilités de Strasbourg*, 35:139-148, (2001).

[30] Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–292, 1979.

[31] Kallenberg, O. *Foundations of Modern Probability and its Applications. Springer-Verlag*, (2002).

[32] Krylov, N.V. A supermartingale characterization of sets of stochastic integrals and applications. *Probab. Theory Relat. Fields,* 123:521–552, 2002.

[33] Krylov, N. V., and Liptser, R. On diffusion approximation with discontinuous coefficients. Stoch. Proc. Appl. 102:235–264, (2002).

[34] Krylov, N. V. On Ito's stochastic integral equations. *Theory Prob. Appl.* 17, 1:20–33, (1972).

[35] Krylov, N.V. *Introduction to Diffusion Processes. Translations of AMS*, 1989.

[36] Krylov, N.V. *Controlled Diffusion Processes. Springer Verlag*, 1980.

[37] Meireles Rodrigues, A.S. *Non-concave and Behavioural Optimal Portfolio Choice Problems* Ph.D. thesis, University of Edinburgh, (2014).

[38] Neveu, J. *Bases mathemátiques du calcul de probabilités. 2me Edition*, *Maison*, 1970.

[39] Prokhorov, Y. Convergence of random processes and limit theorems in probability theory. *Theory of Probability and its applications*, Vol. 1, 2:157–214,1956.

[40] Quiggin, J. A Theory of Anticipated Utility. *Journal of Economic and Behavioral Organization,* 3:323–343, 1982.

[41] Rásonyi, M., and Rodrigues, A. M. Optimal portfolio choice for a behavioural investor in continuous-time markets. *Ann. Finance* 9:291–318, 2013.

[42] Rásonyi, M., and Rodríguez-Villarreal, J.G. Optimal investment under behavioural criteria -a dual approach. Advances in Mathematics of Finance *Banach Center Publ.* 104:167-180, 2015.

[43] Rásonyi, M., and Rodríguez-Villarreal, J.G. Optimal investment under behavioural criteria in incomplete diffusion market models. Available at `http://arxiv.org/pdf/1501.01504.pdf` *submitted*

[44] Rásonyi, M., and Stettner, L., On utility maximization in discrete-time financial market models. *Ann. Appl. Prob.* 15:1367–1395, (2005).

[45] Reichlin, C. Utility maximization with a given pricing measure when the utility is not necessarily concave. *Mathematics and Financial Economics*, 7:531–556, 2013.

[46] Reichlin, C. *Non-concave utility maximization: optimal investment, stability and applications.* Ph.D. thesis, ETH Zürich, 2012. Diss. ETH No. 20749.

[47] Revuz, M., and Yor, M. *Continuous Martingales and Brownian Motion.* Springer-Verlag, 2010

[48] Rockafellar,.R.T. and Wetts, R. *Variational Analysis. Springer-Verlag*, (2009).

[49] Rogers, L.C.G. Equivalent martingale measures and no-arbitrage. *Stochastics and Stochastic Reports* 51:41–49, (1994).

[50] Shiryaev, A. *Probability.* 2nd Edition, 1995, *Springer.*

[51] Stroock, D. *Probability theory, an analytic view.* 2nd Edition, *Cambridge.*

[52] Stroock, D. *Lecture Notes in Stochastic Analysis: Diffusion Theory* 2nd Edition, *London Mathematical Society*

[53] Stroock, D. W., and Varadhan, S.R.S. *Multidimensional Diffusion Processes* Springer, Berlin, 1979.

[54] Stroock D., and Varadhan S.R.S. Diffusion processes with continuous coefficients I, *Comm. Pure Appl. Math.*, 22:345–400, (1969).

[55] Stroock D., and Varadhan S.R.S. Diffusion processes with continuous coefficients II, *Comm. Pure Appl. Math.*, 22:479–530,(1969).

[56] Tudor, C. *Procesos Estocásticos (Stochastic Processes). (in spanish) SMM*, (1999).

[57] Tversky, A., and Kahneman, D. Advances in prospect theory: Cumulative representation of uncertainty, *J. Risk Uncertainty*, 5:297–323, 1992.

[58] Tversky, A., and Kahneman, D., *Extensional vs. Intuitive Reasoning: The conjunction fallacy in probability judgment*, Psychological Review, 90:293–315,(1983).

[59] von Neumann, J., and Morgenstern, O. *Theory of games and economic behavior*, Princeton University Press, 1944.